

TOP SECRET

25X1

25X1

No. Pages 328

15 January 1964

Nº 147

**STAFF FOR THE COMMUNITY
INFORMATION PROCESSING STUDY
(SCIPS)
STAGE I REPORT**

NSA review
completed

THIS DOCUMENT CONTAINS CODE WORD MATERIAL

Volume II

**Findings and Discussion
and
Illustrations**

October 1963

NAVY review
completed.

DIA review(s)
completed.

ARMY, DOS and USAF review(s)
completed.

GROUP 1
Excluded from automatic
downgrading and
declassification

TOP SECRET

25X1

Page Denied

TOP SECRET

25X1

STAFF FOR THE COMMUNITY INFORMATION PROCESSING STUDY
(SCIPS)

STAGE I REPORT

Volume II

FINDINGS AND DISCUSSION
AND
ILLUSTRATIONS

TOP SECRET

25X1

Page Denied

TOP SECRET [REDACTED]

25X1

SCIPS STAGE I REPORT

VOLUME II

FOREWORD

Sections I and II (Summary, Conclusions, and Recommendations) are contained in Volume I. This volume of the Stage I report contains the Findings and Discussion of the study. The principal findings in Section A are statistical and are presented in graphic form -- that is, in pie charts and bar graphs (in Appendix A). The narrative comments in Section A are the result of limited analysis. Section B is a more narrative examination of the data on a problem basis -- problem areas more or less specifically assigned by the terms of reference of the study. Section C is a series of discussions of the "think-piece" type, and Section D is a discussion of the study effort itself. Section B when separated from Section A is classified [REDACTED]

25X1

[REDACTED] Sections C and D when separated from Sections A and B are classified SECRET.

Included also in this volume is Appendix F (Foreign Publications Findings and Analysis), which is a part of the Findings and Discussion resulting from analysis on a "problem" versus a "total" basis. The result is an entity separable and usable at the lower security classification of SECRET.

TOP SECRET [REDACTED]

25X1

Page Denied

~~TOP SECRET~~

25X1

SCIPS STAGE I REPORT

VOLUME II

TABLE OF CONTENTS

	<u>Page</u>
III. Findings and Discussion	1
A. Picture of Community Information Processing	1
Introduction	1
The Intelligence Context	1
The Present Information Processing System	4
1. Organizations, People, and Equipment	5
a. Organizations	5
b. Personnel	7
c. Equipment	8
2. Items	9
a. Definitions	9
b. Item Categories	10
c. Items by Information Types	11
d. Item Descriptions	12
e. Items by Class	15
f. Items by Originator	16
g. Specific Items	17
3. Flows	18
a. General	18
b. Over-All Departmental Flow Patterns	19
c. Interdepartmental Flow Patterns	21
d. "World" Flow Patterns	23
e. Selected Item Flow Patterns	25
4. Processes	27
a. General	27
b. Process Descriptions	27

- v -

~~TOP SECRET~~

25X1

TOP SECRET

	<u>Page</u>
c. Over-All Picture	27
d. Selected Items	28
e. Selected Processes	29
f. Selected Organizations	29
5. Files	29
a. General	29
b. Total Files and Size of Files	30
c. Age of Files	31
d. Present Rates of Growth of Files	32
e. Future Rates of Growth of Files	32
f. Purge Criteria	33
g. Files with Standard Formats	33
h. Manual Versus Machine Files	34
i. Types of Files	34
j. File Content Categories	35
k. Purpose of Files	36
l. Intended Utilization of Files	36
m. Security Classification of Files	36
n. Dissemination Controls on Files	37
o. Consumer Access to Files	38
p. Physical Form of File Items	38
q. Order of Files	39
r. Correlation of Machine Files with:	41
(1) Consumer Access	41
(2) Types of Files	41
(3) Content of Files	41
s. Correlation of the Rates of Growth of Files with:	41
(1) Security Classification	42
(2) Security Classification Versus Form	42
t. Correlation of the Form of Files with Other Factors	43
(1) Security Classification	43
(2) Effective Date	43
(3) Purge Criteria	44
(4) Order of Files	45
(5) Content of Files	46

TOP SECRET

TOP SECRET

	<u>Page</u>
u. Filing Criteria	46
B. Information Processing Problems	47
Introduction	47
1. Indexing	49
a. Index Files	49
b. Indexing Tools	50
c. Element Control	50
d. Common Indexing	50
2. Data Exchange	51
a. General	51
b. The Files Problem	52
3. Report Formatting	54
4. System Utilization	55
5. Biographic Information	57
a. Background and Scope	57
b. Findings	58
(1) Biographic Files	58
(2) Biographic Information Control	64
c. Summary Discussion and Conclusions	66
d. Further Work Required	70
6. Definition of Intelligence	70
C. General Discussion	77
1. The Information Flood	77
2. The "All-Source" Problem	79
3. The All-Source Index	81
4. Security and Information Processing	83
5. State-of-the-IP-Art in Intelligence	86
D. The Stage I Study	91
1. Other Data	91
2. The Study Experience	91

TOP SECRET

~~TOP SECRET~~Appendixes

	<u>Page</u>
A. Illustrations	95
F. Foreign Publications Findings and Analysis [Including 20 Illustrations]	103

Tables

1. Biographic Files, by Agency	59
2. Biographic Files -- Physical Form	62
3. Biographic Files, by Subject Coverage	67

Chart

Categorization of Intelligence and Intelligence Information	73
--	----

~~TOP SECRET~~

TOP SECRETSCIPS STAGE I REPORTVOLUME IIIII. Findings and DiscussionA. Picture of Community Information ProcessingIntroduction

Throughout all the documentation preceding and surrounding the SCIPS Stage I study (see Appendix G, Volume V) is the expressed or implied need for a "picture," a comprehensive, comprehensible presentation of the environment in which information processing systems are operating and management decisions have to be made. A picture, as such, cannot be given, for it exists only in the recipient's mind. What follows, then, in this section is, first, a short, narrative -- subjective, qualitative, and undocumented -- "overview of the intelligence process" and then a presentation of literal and graphic symbols, representing quantitatively the real world (within the study scope previously described), which is intended to facilitate the reader's gaining a factual comprehensive "picture" of information processing in the US intelligence community.

The Intelligence Context

The staggering size, number, and variety of information processing (IP) operations (systems) barely disclosed in the Stage I study, together with the known related research and development (R & D) activities, call for a redescription of the intelligence context -- that is, the total system, within which the present IP operations of concern here are conducted.

The intelligence production process is coupled with a versatile, dynamic collection effort that began with the acquisition after World War II/ This beginning has been followed by a series of increasingly sophisticated collection techniques. The cumulative result of these efforts is a large daily flow of information and the existence in files of a tremendously large, dynamic data base of potentially useful information dating in some categories from 1940. The increasing amount of information both incoming from collection and in files for retrieval and retrospective exploitation presents the intelligence analyst with an increasingly difficult task of bringing to bear all the potentially significant information either on his current or on ad hoc functional responsibilities.

25X1

TOP SECRET

TOP SECRET

25X1

Requests for collection by field activities usually are restricted to "gap filling" and supposedly are utilized only after available information has been exploited. Rather elaborate techniques are employed by collection managers to avoid levying duplicate requests for collection, but little organized effort is made to avoid collection of information already on file. It is more often a matter of subject-area priorities competing for limited collection-exploitation resources.

As a generalization, field collection is conducted in accordance with both general and specific requirements developed by analysts in functional and area research components throughout the community. By far the greatest volume of information is collected in response to general requirements (assigning a number does not make the requirement specific) and flows continuously into central processing areas. This flow is screened and routed to analysts in accordance with their functional areas of responsibility, sharpened somewhat by subject-area statements of interest (hereafter termed "reading requirements" to distinguish from "collection requirements").

Most collected information serves a number of purposes in a variety of time frames. For example, a given item of information may reach "warning evaluators and decision-makers minutes after its acquisition. At the same time, this same item of information will start from its acquisition to numerous recipients for basic, operational, command, interdepartmental, and national intelligence production. The same event or situation may be reported by a number of sources and start forward from its several acquisition points as described above.

The amount of unique information incoming from collection and on file for retrospective exploitation imposes a tremendous burden that may preclude effective exploitation. In some cases, only the most cursory examination can be accorded many items as volumes increase and backlogs build up at analysis or processing points. Receipt of duplicate reporting at all echelons in all organizations may reduce further the effective utilization of available analytic resources. Although there are problem areas of intelligence where the paucity, not the deluge, of reporting is the problem, the latter is more common and indeed may be more closely related to the former than is apparent. Literally hundreds of thousands of copies of items move daily in the community.

Analysts' reading requirements are general enough to reserve to themselves the selection of items of marginal interest. In addition, many analysts maintain that they need the complete document in order to evaluate adequately the items of specific interest contained therein. These conditions tend to transfer effective screening from the dissemination services to the individual analyst. This screening that is reserved by the analyst has impact only on his immediate

- 2 -

TOP SECRET

25X1

TOP SECRET

disseminator and not on the many prior judgments made as information moves from collection to analysis.

Experienced analysts have come to recognize that different functional points of view, when applied to the same body of information, tend to surface items of significant interest which might otherwise have been overlooked. Consequently, many analysts rely on contact with their counterparts in their own and other agencies for the purpose of surfacing significant items that they may have missed.

Having established his own, sometimes unique, information acquisition subsystem by the two foregoing principal methods, the intelligence analyst is ready to estimate the reality of foreign posture and purpose as reflected in the information available to him. This estimation may involve continuing identification and association of activities, events, and situations reported from any of a number of acquisition methods.

As a new item of information is identified, for example, it is compared with the current estimates that the analyst is maintaining. The item is evaluated in light of similar or related information that was received and retained in the past. On the basis of the new information, the analyst may revise his estimate, redefine or describe a physical characteristic, or correct an installation location. He may determine that the new information represents normal activity or is indicative of a developing capability or trend portending new enemy posture or purpose. This sequence may be termed the continuing "analysis-update" process. There are numerous advantages to this process. Because it is based on assigned areas of analytic responsibility, it usually is consistent with the mission and functions of the parent agency and results in a proportionate but even less distinct division of functional processing responsibilities among agencies.

Information continually is filtered and incorporated into his estimates by the analyst. The analyst can adjust to a degree the volume of information flowing to him by raising or lowering the specificity of his standing "reading requests." In periods where maximum output of estimates is being stressed, the "analysis-update" process can be slowed. Conversely, when output requirements recede, the "analysis-update" process can be expanded.

There are, however, other characteristics of the continuing "analysis-update" process. In practice, heavy reliance is placed on each analyst to identify and exploit information pertinent to his area of assigned responsibility. Of more significance to analytic support, however, is the fact that the "analysis-update" process does not respond readily to unanticipated problems. When such problems arise,

TOP SECRET

TOP SECRET

the location and identification of pertinent information in the analyst's working files may be extremely difficult, if not impossible, to accomplish, because the analyst did not retain, identify, or index that information. Time-consuming item-by-item inspection in fact may be required if these working files are to be used.

The search must then be shifted to files established to provide ad hoc response to general problem areas. There are hundreds of these files being maintained throughout the community. Most of these files are products of evolutionary development and respond to the purposes set forth from time to time by the holding organization. Access to these holdings may be direct through the "file" order in which the items are disposed or may be enhanced through the creation of bibliographic references, abstracts, extracts and/or indexes that may exist in the form of handwritten or typewritten cards in a variety of sizes. The files also may be placed in their entirety or in part on punched paper tape, reel film, film segments, or film chips or on magnetic tapes or discs. The written or typewritten cards themselves may be reduced to microfilm in a variety of forms. Machine-coded references may be combined with the document image and thus constitute a single unit, or the item and its index data may be maintained separately. It is conjectured that virtually all items of information find their way into at least one file in the community. Some items also may find their way into most files, but it is a certainty that there is no single file or coordinated group of files in which "all" valid items may be found. The analyst with a new problem or a new aspect of an old problem is in trouble if he wishes to examine or reexamine all pertinent information on a given subject or area. Thrown on his own resources, he will peruse as many or as few community files as he deems appropriate to the problem or as he has time to examine. Availability of pertinent information in this context becomes a function of an analyst's experience, insight, and diligence as well as of the accessibility, content, organization, and retrieval characteristics of the various files -- rather than a function of whether or not the information ever has been collected and disseminated. On the other hand, whether the analyst is conscientious or not, he can be assured that no one else can afford sufficient time to demonstrate conclusively the degree of adequacy or inadequacy of his file search. Or, expressed another way, the present system places an awful burden of conscience on the conscientious analyst-estimator.

The Present Information Processing System

However nebulous, abstract, or disjointed the "community IP system" might be, its principal constituent parts are hard and real and should be nameable. The "system" consists of organizations including processing people and equipment, items of intelligence or intelligence information in some physical form, the movements or flows of

- 4 -

TOP SECRET

TOP SECRET

items between organizations, the processes performed by the organizations on the items, and the repositories or files of items. There follows here a section on each of these five principal parts (1 through 5, below). Elementary graphics (to be found in Appendix A) are used extensively in the presentation of this study, with a minimum of corrective, expository, or qualifying textual material. This method has three facets: (1) what little text there is in the way of legends and labels on the graphics requires attention, (2) the reader has to exercise more initiative to "receive the message," and (3) the reader is not constrained in the messages that he receives by receiving only the author's view of the data.

Each section starts with generalizations or summary statistics and proceeds somewhat hierarchically down to constituent elements or parameters and then presents some correlation of these parameters. In drawing conclusions from these "pictures," the reader is advised (1) to bear in mind the scope of the study, (2) to consider order of magnitude rather than specific quantities, and (3) to look for proportional relationships as well as absolutes.

1. Organizations, People, and Equipment

a. Organizations

It is estimated that there are some 1,500 to 2,000 addressable organizational units involved in one way or another with intelligence IP. There are in the SCIPS data base (including obsolete, redundant, and incorrect information) 2,985 different "organizations" sending or receiving intelligence items. ³²/* In the SCIPS data base an "organization" can vary from a one-man attaché post to a 1,500-man reconnaissance squadron. Of all the organizations involved, some 200 to 250 are considered to be of primary concern to a study of the community IP system.

Exhibit u, Appendix G,** lists in abbreviated form these principal "organizations" ("orgs") in three groups. The 130 organizations listed in Groups I and II might be considered the major constituents of the community IP system, but those in Group III must be considered significant in interaction with the others. Exhibit g (3) is the organization code developed for the purposes of the SCIPS study and gives the full name and departmental subordination of all organizations that were expected to occur in a study of community information processing. In Exhibit u, the organizations in Group I

* See the list of references in Appendix G, Volume V, and in Volume I.

** See Appendix G, Volume V.

TOP SECRET

TOP SECRET

25X1

are those organizations studied during Stage I. Exhibit h provides the full titles for those organizations and indicates the varying methods and depth of study applied in Stage I. As discussed in the introduction to this study, the information obtained from these organizations provided much item flow information on other organizations; and, in general, the data base from which the following picture is derived might be considered to represent some 35 percent of the total. In terms of items and flows, the proportion perhaps would be 50 percent but would be lower in terms of people, equipment, and processes. In addition to this immediate section, the sections following on "items" and "flows" contain part of the "organizations" picture.

The distribution by department of the some 60 organizations surveyed is shown in Figure III a.* Thirteen of the organizations surveyed are treated separately in Appendix H, Volume VI. Of the other organizations, 47 were treated by direct methods (that is, some information on organization, personnel, and equipment was obtained) and are distributed by department as shown in Figure III b. Within these "organizations," some 300 "processing activities" were identified and described. Establishing a "processing activity" was at the discretion of the surveyor, but they are in general the next level below "organization" and represent points in or aspects of the processing system. In the latter sense they represent more adequately the structure and content of both the data base and the real world than does the formal organizational structure. The "processing activities" also represent a structure much less perishable than the formal organizational structure. A list of the processing activities is given in Appendix C, Volume III, and their numerical distribution among major departments is shown in Figure III c.

Figures III b and c and Appendix C, Volume III, imply that the study has a strong bias toward headquarters and CIA and away from Navy in comparison with proportions in the real world of community information processing. The bias toward headquarters was by design; the other bias was a result of staffing and study circumstances.

Two other observations can be made. The first is that the number of departmental major components surveyed is high, and the second is that, of the some 60 organizations surveyed by one method or another, only 15 are under the direct operational control of a CODIB member.

The average number of organizations per department was about 10. The average number of processing activities per organization was about 6, and thus about 60 processing activities exist per department.

* See the illustrations in Appendix A.

TOP SECRET

25X1

TOP SECRET**b. Personnel**

The SCIPS Stage I study was anything but a manpower utilization study. Although the number and type of people involved is certainly a part of the information processing "picture," it was believed that existing studies on this aspect were better than could be done within the framework established for Stage I. Thus the recording of personnel done in the Stage I survey was done as a check point on magnitude of processing rather than on utilization of personnel.

In those organizations studied (see Exhibit h, Volume V, Appendix G), only the number of people involved in the processes being surveyed and described was recorded. Sometimes an estimate of functional equivalence was used to determine the "number of personnel." Thus the study figures on quantities of personnel will not necessarily agree with official figures for the named organizations.

The some 5,000 personnel recorded during the survey were distributed by department and between "clerical" and "professional" as given in Figure III d. In the aggregate, there were about 17 people per processing activity and almost a 1 to 1 ratio of clerical to professional personnel. There are, however, some extremes in the clerical to professional ratio -- 1 to 3 in the Army versus 14 to 1 in CREF/NSA. The survey instructions were to the effect that the requirement for a person holding a college degree or a commissioned officer distinguished for study purposes between "professional" and "clerical." It would be expected that a higher clerical to professional ratio would be found in the "processing" activities than in "intelligence analysis" activities but that the ratio would be lower than in logistic or administrative support activities. Figure III e is an attempt to correlate the clerical versus the professional factor with actual processes being performed. Because the study method did not attempt to record the type of personnel with each process but only with groups of processes, only the extremes in Figure III e could have any validity. A special study would have to be made to be conclusive on personnel utilization. On an individual process basis, Figure III e does show a higher professional to clerical ratio on such processes as "abstracting," "dictating," and "translating," as would be expected. It is observed then that these processes constitute a smaller part of the present community IP than the processes done by clerks. Surprisingly, activities involving the "filing" process have an almost 2 to 1 ratio of professional to clerical, whereas activities involving "indexing" have only a 1.3 to 1 ratio of professional to clerical. Again, the detailed personnel information of Figure III e is very highly suspect and serves primarily to show a potential area for special study.

- 7 -

TOP SECRET

~~TOP SECRET~~

25X1

c. Equipment

A comprehensive inventory of IP equipment was not an objective of the Stage I survey. Stage I was concerned with the identification of what processes were being done rather than how the processing was done. Nevertheless, the association of equipment with the processing activities being surveyed serves to identify processes and are part of the "picture." Only that equipment was recorded that actually was used in the processes surveyed. Planned equipment, ownerships, and proportionate time of utilization were not survey factors in Stage I. Within the scope of the study the equipment information that was recorded makes an interesting picture in itself, as shown in Figure III f. An equipment code (see Exhibit v, Appendix G, Volume V) was developed after the field survey on the basis of actual equipment noted and was applied to categorize and group "like" equipment. Figures III f, g, and h add unequal equipment in portraying some 3,000 pieces of equipment by general category and by agency. The full listing of surveyed equipment (except that covered in Appendix H, Volume VI), is given in SCIPS Data Catalog R2 (Appendix E, Volume III). The quantity of reproduction and duplicating equipment is worthy of note (see Figure III f) as well as its pattern of distribution among departments (see Figure III h), which is very similar to "studied-activities" distribution in Figure III c. The commonness of reproduction equipment is noted -- that is, 27 out of 45 "orgs" studied had reproduction equipment (Figure III i). Many of the "orgs" were subcomponents of a major component.

Figure III h contrasts the varying popularity of the three important categories of equipment, by agency. The distribution of electronic data processing (EDP) and electric accounting machines (EAM) equipment by department in Figure III h probably would not vary greatly if the study were broadened. Most of the Army and AFIC equipment shown in Figure III h is now organizationally in the ADP center of the Defense Intelligence Agency (DIA). The low count of EAM and EDP equipment in NSA is relatively accurate for the activities studied in Stage I (that is, CREF).

Figure III i presents quantities of selected items of equipment of particular interest. In addition, Figure III i shows how many of the different surveyed organizations used each particular type of equipment. The commonness of the typewriter should not be but always is a little surprising. A greater surprise, in that the survey did not particularly hit the originating or generating organizations, is the number of flexowriters (77) in different organizations (17). Another surprise is the number of teletypewriters (51 in 7 organizations), for communications centers were not studied as such. A minimum of 150 pieces of EDP equipment in some 20 EDP installations were found to be

~~TOP SECRET~~

25X1

TOP SECRET

used by 12 different organizations surveyed. Of these 20 EDP systems, there was considerable variety (12 different systems and 6 different manufacturers). In terms of pieces the EDP systems are approaching or have exceeded the EAM systems -- undoubtedly at a considerably higher average cost per piece. At the present state-of-the-art the EDP equipment is in competition more with EAM equipment than with microfilm equipment, and this may remain the situation for some years.

2. Items

a. Definitions

One of the most difficult pieces of the "picture" to comprehend, much less to tie down, is the objects on which the processes are performed or which result from a process by the personnel and equipment in all those organizations, that which flows between organizations, or those things that are put into all the files. These things are called "items." Thus the term item applies not only to documents and reports but also to newspapers, magazines, periodicals, monographs, books, forms, memos, manuals, maps, charts, punched cards, magnetic tape reels, microfilm chips, photographs, "teletypes," messages, cables, and even some not so tangible things such as "verbal requests." The definition and concept of an item as used in the field survey during which specific items were initially identified are given in Exhibit g (1), Appendix G, Volume V. In general, an item is a series or group of individual physical pieces or issuances. The survey method (that is, simultaneous survey in different organizations by different surveyors) resulted in duplicative but dissimilar identification of items. The same item would be given a different identification number (and sometimes a different name) in different organizations. These unique names were called "local items," and some 32,000 of them were identified during the field survey. ^{33/} During analysis, and with the aid of computer print-outs, local items were reexamined and compared for similarity and identity and were reidentified as unique "SCIPS items" (see Exhibit s, Appendix G, Volume V). This analytical process resulted in the identification of some 14,000 "SCIPS items." ^{34/} Although this fact would indicate that the survey of additional "orgs" would result in the addition of proportionately that many more different items in the picture, study experience is that the some 14,000 SCIPS items are the bulk, perhaps 70 percent, of the total picture. Missing items are primarily local products of processing that do not flow between organizations and field items receiving lateral dissemination only. Although the over-all ratio of "local items" to "SCIPS items" is 2.3 to 1, a breakdown of this shows that in the activities surveyed in Stage I the same foreign publication item was found on the average in 1.7 different organizations, whereas the non-foreign-publications item was found in 3.6 different organizations. Survey

TOP SECRET

TOP SECRET

25X1

bias on foreign publications was toward a lower ratio than the real world, and on non-foreign-publications survey factors made for a lower ratio, but SCIPS item numbering aggregations made for a higher ratio. In both cases the ratio is lower than was anticipated.

b. Item Categories

With the variety of composition in mind the ambiguous term items can be used, and the categories or classifications of items can be considered for the sake of comprehension. In this further comparison between items it must always be borne in mind that different items are not equivalent in many ways. For example, index transcript sheets may be issued once in one copy and consist of one page, whereas a foreign technical journal may be received in 12 issues per year in 5 copies of 150 pages each, or an Air Force information report (IR) may have 25,000 issues per year in 50 copies that average 2 pages each, but each of these three "groups" is a "SCIPS item."

During the field survey, two different classifications of items were made. The first of these was called "Category." These categories were as follows: SB, substantive; SA, substantive aid; RX, requests; and NS, nonsubstantive. The category "nonsubstantive" was a manner of controlling the scope of the Stage I study rather than a literal meaning. Nonsubstantive items were administrative reports, field collection requirements, and others not of direct concern to the scope of the Stage I study but were mixed in with substantive items in the real world. An example of the basic difference between SB and SA was the contrast between an Air Force IR and an index to Air Force IR's or a manual on how to index IR's. In practice the distinction is not so clear-cut. Figure III j is a breakdown of the SCIPS items among these major "Categories." Any depiction such as Figure III j requires immediate further consideration of the overwhelming category -- in this case the SB, or substantive category. Figure III k identifies by far the largest single constituent of the substantive category of items. The some 10,000 SCIPS items of open literature foreign publication are not a monolithic group either, in that they include newspapers, magazines, and monographs both in foreign languages and in English. Although they occupy some 75 percent of the community item picture, they receive neither typical distribution nor typical processing. For these very same reasons and for purposes of comprehension, the foreign publication items, unless specifically noted, are broken out from the rest of this section and are treated separately in Appendix F.

TOP SECRET

25X1

TOP SECRET [REDACTED]

25X1

Reconsideration of the SCIPS items (minus foreign publications) by category results in Figure III l. On the basis of items the proportions shown in Figure III l are believed to be a reasonably true reflection of the real world with three qualifications, as follows: (1) nonsubstantive items were intentionally minimized during the study; (2) the number of request items must occupy a larger share in the activities that were studied (see the discussion in III, B, 4); and (3) reclassification probably would shift some items from the "substantive" to the "substantive aid" category. In terms of volume and commonness the "substantive" category items usually are of multiple issue and flow between organizations. As a group, many but not all of the "substantive aid" items are local intermediate products in single copy.

c. Items by Information Types

Even before the field survey system was designed, the staff made several attempts to categorize organizations, activities, processes, and items. One of the more useful categorizations was what was called the "worlds," which comprise the intelligence "universe." The establishment of this system was a way of breaking down and grouping identifiable parts of the whole picture. Sixteen "worlds" were discerned and given informal unofficial labels. These "worlds" were used to categorize items during the survey and are portrayed graphically in Figure III m. Around the "world" of reference intelligence, with which the SCIPS effort was primarily concerned in Stage I, are the other two derivative worlds: analyzed intelligence and initial intelligence. These three worlds are set in a field of blue comprised of 13 primary worlds, as follows: Seismic, Radar, Acoustical, Communications, [REDACTED] Intelligence, Photography, Human Observation, Foreign Broadcasts, Foreign Publications, Mapping, Foreign Materials, and [REDACTED]. These 13 primary worlds are mostly readily identifiable with specific sources or collection techniques, but many intelligence items are of more than one world. An item categorized in one of the primary worlds generally retains its identity. When the primary items enter current intelligence, analytic production, or central reference, they are processed with other items and lose their original individual identity -- that is, one-for-one identification is no longer possible, but composite or new items are generated that have their own identity and in turn enter the processing system. The worlds thus are a means of categorizing "things" -- in this case, SCIPS items. Figure III n (1) is a distribution by "Info Type" or "world" of the SCIPS items. Figure III n (2) is the same except that Foreign Publications are excluded. Because the SCIPS item concept had some variation between worlds, a picture based on "issues" rather than items might be a truer representation of the real system. Figure III n (3) is a picture based on number of issues per year rather than number of items, as are Figures III n (1) and (2). Data on the number of issues were not collected on all

25X1

25X1

TOP SECRET [REDACTED]

25X1

TOP SECRET

25X1

items during the survey. The quantities shown in Figure III n (3) are based on survey data, whereas the quantities in parentheses are estimates based on a few known individual items and on a great deal of "judgment."

The proportions in Figure III n (3) between worlds are believed to be representative of the unique data base in the community in terms of item issues. A similar distribution on the basis of "copies" rather than issues would be a "total paper" picture. Although this was not done during the study, it is known that more copies per issue are generated for items in the COMINT, HUMINT, FIBINT, INTINT, and ANLINT worlds than in the PUBINT, REFINT, and PHOTINT worlds. The "open" worlds of PUBINT and FIBINT comprise essentially unclassified information and are at least relatively cheap to collect on a basis of cost per item. The world of HUMINT is comprised of items generally classified CONFIDENTIAL or SECRET that supposedly are of high pertinency because of the controlled reporting against requirements needs by controlled collectors. The world of COMINT is one of high cost, uniqueness, and high sensitivity. The world of PHOTINT, likewise of high sensitivity, has the advantage of validity and the disadvantage of high-cost exploitation. These variable features of the principal primary worlds result in varying IP resource allocation and techniques.

d. Item Descriptions

Aside from identifying, categorizing, and counting "items," it would be desirable and necessary for analysis to describe the physical and informational content characteristics of the items. These "physical" and "content" characteristics were determined and disposed on the Item Description Sheet (IDS) Survey form [see Exhibit g (4), Appendix G, Volume V].

During the course of the study, there were some 400 SCIPS items that were considered to be intermediate products -- that is, items generated as a result of or for the purpose of processing other items. For example, transcript sheets (from which punched cards are made to serve as an index to a document file) are considered to be intermediate products. These kinds of items usually do not flow between organizations. Although a correlation has not been calculated, it is assumed that nearly all of the 400 intermediate products are in the "Substantive Aid" category and the "Reference Intelligence" world. Because of study constraints (see Section IX, B, of Appendix G, Volume V), and because these were products of the organizations surveyed rather than of the "user" organizations, descriptive information was obtained on a much higher proportion of the intermediate products or "Substantive Aid" category of items than for the "Substantive" category items. Of some 400 SCIPS items that are intermediate products,

- 12 -

TOP SECRET

25X1

TOP SECRET

25X1

Item Description Sheets were completed during the field survey on three-fourths of them but on less than 400 of the other 3,000 SCIPS items. With these biases in mind, including the smallness of the sample (800 out of 14,000 items), the following descriptive information was obtained:

- (1) Type of Information. The distribution of SCIPS items by type of information given in section 2C and Figures III n (1) and (2) was determined for all SCIPS items during the analysis. During the survey phase, both primary and secondary information types were recorded for some 655 of the 800 SCIPS items with IDS's. Of the described items, the vast majority were ranked first in the REFINT world. HUMINT, COMINT, and PUBINT dominated the second rank, indicating that most of the REFINT items are derived from those three worlds plus the PHOTINT and MAPINT worlds.
- (2) Language. The language of the item was recorded for 815 SCIPS items. As shown in Figure III n (4), nearly three-fourths of the items were in English and involved no translation or transliteration from a foreign language. Description sheets were not completed on the 10,000 foreign publications, most of which are of course in a foreign language; so the surprise in Figure III n (4) is that one-fourth of the other items did involve a foreign language.
- (3) Subject Content. The distribution of described items by broad subject categories is shown in Figure III n (5). Of the 815 SCIPS items, more than one-half of them were described as containing information on all subjects (SU), and nearly another one-third of them were assigned a single subject code. It is in the small remaining portion that the multiple code assignment occurred. Those 376 that were assigned the all subjects (SU) code can presumably be added to each of the named subject categories. Military, S & T, Installations, and Biographic comprise the bulk of the items. Because of the organizations studied and, therefore, the sample bias in items described, it is suspected that Figure III n (5) is more representative of items that receive a lot of processing rather than a picture of all items originated.

- 13 -

TOP SECRET

25X1

TOP SECRET

This in itself causes interesting speculation, but a broader sample should be acquired before a conclusive answer is derived.

- (4) Area Content. Figure III n (6) shows the distribution of the 815 described items by broad geographic areas representing item content. There were about 1.1 area codes assigned per item. All of the 423 "all area" items presumably could be added to each of the individual area code categories. Only some 11 percent of the items did not concern a Communist area. Again, it is suspected that this is more representative of item processing by central reference facilities than of the intelligence items generated in the community. If valid, this ratio might relate to rates of reference system utilization.
- (5) Physical Description. In addition to the substantive content of items, the physical form and carrier of the information item has equally determining influence on the processing of the item or is the result of a given technique or process. When the same item of information exists in more than one physical form or carrier, there is a most difficult problem of item identification. Any change in form for the same item involves some processing -- sometimes through laborious manpower and sometimes more automatically. In this sense the physical form of items is a measure of the result of processing and, therefore, of utmost importance to an analysis of the community "IP System." Unfortunately the sample of items so described in Stage I is small -- approximately 800 SCIPS items. The sample also is biased, but in this case desirably so, toward the products of "central" information processing activities. Because of varying levels of "item" definition and uncontrolled use of units of measure (number of documents versus reels), considerably more analysis or resurvey is required to derive definitive item volume figures ("issues per year"). Even so, the physical description of some 300 different "SCIPS items" warrants consideration. The development of a physical

TOP SECRET

TOP SECRET

descriptive system for items is of considerable value in itself. The form of an item is expressed in terms of three characteristics, as shown in Figure III n (7). As can be readily discerned, some 80 percent of the items are typed or printed text, readable only visually. The bulk of the machine readable items are digital and mostly punched cards. The distinction between "typed" and "printed" was not too valid. The number of items that were both visually and machine readable is small, about 22 percent, and are mostly interpreted punched cards. The "carrier" of an item is described by two characteristics, "material" and "shape." The distribution of the some 800 described SCIPS items by these two characteristics is shown in Figure III n (8). The carrier is not necessarily "predictable" from the "form." Although forms implying advanced mechanizations (reel, disc, chip, and strip) made some showing, the preponderance is paper sheets and cards. About half of the "card" items are punched cards. So again, on an item basis, a maximum of 20 percent mechanization is indicated.

e. Items by Class

During the process of surveying the determination of what was a local item was difficult and necessarily inconsistent. In the post-survey period of analysis when the local items were reexamined, identified, and given a SCIPS item number, they also were "classed" by the last position of the SCIPS item number assigned to them (see Exhibit s, Appendix G, Volume V). The some 10,000 foreign publication items were, of course, primary items (Class A). The other 3,800 SCIPS items that were classified are distributed between classes in Figure III p. The items in the last two classes, "Item Groups" and "Super Items," represent failures in survey or management record keeping. The second class, "Sub Items," represents both an unresolved conceptual problem and a survey problem as well as a real world system complexity. Within the first and most important (quantitatively) class is found the bulk, 80 percent of which are "A's" -- that is, "primary item" in their original or most common and complete form. Even so, there are 408 items in Class "B" -- that is, complete but not in the original or common form -- and 235 items in Class "C" -- a piece or part of an original item. Classes B and C imply that some transforming process has been performed on the original item. Considering the activities surveyed, it is somewhat surprising to find so many of the items non-transformed. Although a picture based on the number of issues versus

TOP SECRET

TOP SECRET

25X1

"items" might be quite different, there is no particular reason to suspect so.

f. Items by Originator

Another obvious way to categorize, group, or classify "items" is by the organizations generating them. Aside from categorization purposes, identifying the originators of items is an integral part of both the information processing "picture" and a picture of the community itself. The difficulty in performing this task involves determining the level of the originating organization that should be identified for comparison purposes, but identification of the originator of a particular item often can develop into quite a research job. Identifying the originator of items was first attempted during the field survey. Unfortunately, most organizations did not know what specific items they received, much less who originated them. In the post-survey analysis period and again with the aid of computer print-outs, the second attempt to identify originators of items was made as part of the SCIPS item numbering exercise (see Exhibit s, Appendix G, Volume V). After the SCIPS item numbers were assigned to all items, an ordering of the Item Data Catalogs by SCIPS Item Number will group the items by originating organization. 34/ Figures III q (0-12) depict the numerical distribution of SCIPS items by originating department and major components. Again the scope of the study and its limitations should be borne in mind. Also in viewing Figures III q (0-12), it must be remembered that "items" can vary from 1 to 20,000 or more issues per year. Unfortunately, during Stage I the number of issues originated could not be determined on a great number of items, which would have presented a "volume generated" picture. Figure III q (0) shows the distribution of items by department. It is believed that this is a fair representation of the real world with two exceptions, as follows: (1) the foreign publication items, all 10,000 of them, are not included in the charts (the originator in this case is of course the multitudinous foreign publishers of open literature, and (2) the study bias (the organizations surveyed in Stage I) has minimized those local and lateral items in field intelligence components of all departments. It is not believed that the latter bias materially affects the proportions in Figure III q (0), but it might affect some of the backup charts of Figure III q (1-12), which are detailed charts for each department that show which major components are originators of items.

In Figure III q (0), in addition to noting that some 40 percent of the items are originated in CIA and the Air Force, it also is of interest that aside from foreign open-literature publications (10,000 items), there is an appreciable number of items (about 800) processed in the community which are not originated by a community component. On the backup Figures III q (1-12) the following pieces of the picture stand out:

- 16 -

TOP SECRET

25X1

TOP SECRET [REDACTED]

25X1

- (1) The number of major components generating intelligence items is very large (approximately 100).
- (2) Some two dozen major components must be named to cover even the bulk of the items.
- (3) In addition to the USIB principals' headquarters components, the big originators of intelligence items are foreign publishers, US military commands, the Department of Commerce, [REDACTED]
[REDACTED]

25X1
25X1g. Specific Items

In addition to defining and categorizing, there is another way to depict the item part of the community picture, and that is to name the individual items. Although this has been done in the SCIPS Data Catalogs ("T" series), one of which (T 9) is Appendix B, it is not possible to retain as a composite picture some 14,000 titles or 4,000 or even 1,000 (nonforeign publications). By the same token, there was not time during Stage I to analyze all the items. The senior SCIPS analysts as a group made a selection of "significant" items for further analysis. The basis for the designation "significance" was never defined other than to have some representation for most "worlds" and originators. The selected items are listed below as an aid in lending some specificity to the item picture:

<u>Item</u>	<u>Originator</u>
(1) Target Data Inventory	ACSI/Air Force
(2) Ground Order of Battle Summary	ACSI/Army
(3) Port Studies	ONI
(4) Sino-Soviet Air Order of Battle	ACSI/Air Force
[REDACTED]	
(6) Consolidated Translation Survey	FDD/CIA
(7) Airgram	State
(8) Summary Translations (as a series)	FDD/CIA
(9) Scientific Information Reports	FDD/CIA
(10) Combat Equipment Technical Intelligence Bulletin	ACSI/Army
(11) Intelligence Bulletin	USARPAC
(12) Quarterly Intelligence Review	USAREUR
(13) ELINT Review	EUCOM

25X1

TOP SECRET [REDACTED]

25X1

TOP SECRET

25X1

Item	Originator
(14) Organizations and Personalities Series	NSA/Headquarters
(15) Headquarters "Translation" Series	NSA/Headquarters
(16) Field "Translation" Series	NSA/Field
(17) Field "Report" Series	NSA/Field
(18) Intelligence Bulletin	DIA
(19) Foreign Broadcast Information Daily	FBID/CIA
(20) Radio Propaganda Report	FBID/CIA
(21) Current Intelligence Weekly Summary	OCI/CIA
(22) National Intelligence Survey	USIB
(23) Economic Reports	ORR/CIA
(24) [REDACTED]	
(25) National Intelligence Estimate	ONE/CIA
(26) Scientific Intelligence Digest	OSI/CIA
(27) Intelligence Publications Index	OCR/CIA
(28) Special Register Title Index	OCR/CIA
(29) Biographic Research Aids	OCR/CIA
(30) Sino-Soviet Affairs	INR/State
(31) Airfield and Seaplane Stations of the World	Air Force
(32) Missiles	FTD/Air Force
(33) [REDACTED]	Air Force
(34) [REDACTED]	JCS/Commands
(35) Information Report	Army
(36) Information Report	CIA
(37) Doklady Akademii Nauk	USSR
(38) Pravda	USSR

3. Flows

a. General

With a picture of the organizations and items in mind, a picture of the interactions between those organizations in terms of the items is now in order. On the basis of this picture, it will be asked in the next section what processing is done on those items by which organizations. Two factors complicate the obtaining of a good picture of item flow -- item identification and organizational level.

TOP SECRET

25X1

TOP SECRET

25X1

Variances in item identification during the field survey could be and were reconciled to a reasonable extent, during analysis. However, where items were aggregated into unidentified "groups," no amount of analysis sufficed. For flow purposes, organizational level proved to be the worse problem. The organizations surveyed usually had records on what items were sent out and to whom, but they often kept no records on individual items received and hardly ever had good records on specific sending organizations in other agencies. This fact is in part due to intervening "mail rooms" that deal in "packages," not "items." As a result, the data base fallaciously shows the same item being sent to the same organization but at two different organization levels. Unfortunately this also actually happens, but there is no way to distinguish in the data base which is which. With specific items identified beforehand and by going to all output points in the community, one probably would obtain a fairly good picture of item flow. This practice was not followed in the Stage I survey. Again bearing in mind the scope of the study and looking for patterns and gross proportions rather than absolute quantities, some valid impressions of the community picture of item flow can be obtained from the Stage I data base.

b. Over-All Departmental Flow Patterns

Figures III r (1-17) depict, by agency, the proportions of internal and external flow. Each chart expresses the flow pattern in terms of "Items," "Issues," and total "Copies." Because of study biases, item definition, or accuracy of reported information, one or the other of the terms may be more accurate than the others in a given case. Only somewhat additive types of items were considered in these flow charts -- that is, "documents," "cards," "mats," "cables," "folders," "forms," "lists," "maps," "photos," "issues," "reports," and "messages" were counted but not "magnetic-type reels," "discs," "decks," "calls," "files," "lines," "strips," and "words." Thus this is sort of "document" flow, meaning a downward bias in terms of all items that flow between orgs. On the contrary, there is an upward bias of even greater magnitude in the absolute volume figures for "items" and "issues" because of organizational level. For example, an item sent by NSA to both SAC and FTD would be counted as two "items" going from NSA to the Air Force. Thus the "item" and "issue" figures are really "item-org" and "issue-org" occurrences. The figure for "total copies," however, is more valid. Despite these general limitations and other specific deficiencies in the data base, there are some consistency and rationale in the charts that warrant attention. Some of these follow as observations:

(1) Figures III r (1-6) portray internal versus external flow. The general pattern is one of considerably more intraagency activity than interagency activity, especially in the case of CIA. Because of the bias of

TOP SECRET

25X1

TOP SECRET []

the coverage of the Stage I study, further study probably would increase slightly the proportion of internal versus external activity in all agencies, particularly in headquarters. Although further study of field activities would reveal additional interagency flows, it is doubted that it would change the over-all picture. Thus Figure III r (1) perhaps is a reasonable exposition of intra-departmental versus interdepartmental intelligence item flow. If so, then for every interagency relationship, there are four or five intraagency relationships. Although quite believable and logical, this fact is somewhat sobering to an interagency study such as this.

(2) Except for CIA, there is startling consistency in the total number of issues and number of copies involved in each agency flow picture; about 2 million to 5 million "issues" per year and about 8 million to 9 million copies (except for Navy at 15 million). The exceptionally large figures for CIA are explainable only in part by the study bias, and extending this same study coverage bias to the other department figures leads to the conjecture that full coverage would reflect higher flow volumes for Navy than for the other departments. The coverage bias toward CIA was not by a factor of five; thus one could conclude that there is more internal paper movement (implying either better coordination or worse duplication, or both, or some other factor not as yet discerned) in CIA than in other departments. Comparison among agencies on the number of items received is believed to be quite invalid because of the bias on the organizational level.

(3) There is a consistent ratio between departments of "copies per issue" of two to four copies per issue. This ratio is both surprisingly consistent and somewhat low. The problem of organizational level has a much more inflationary effect on the "issue" figures than on the "copies"; thus the true ratio of copies per issue would be higher than the 2.3 indicated.

(4) The total "item-org" occurrences for all six agencies equals about 78,000, which shows a minimum five-fold inflationary factor above the some 13,000 to 14,000 unique SCIPS items. The total "issue-org" occurrences equal about 38 million. The assumption of the same inflationary factor reduces the total to about 500 issues per item, or some 7 million issues per year of the

TOP SECRET []

TOP SECRET

14,000 SCIPS items. This is a rather astounding but believable figure. If a cross-check made by researching each item to determine how many issues originated were to verify the figure of 7 million, then we would have a fair idea of the size of the community's unique data base -- that is, perhaps some 10 million item issuances per year.

(5) Reaching for an "all-all" figure, we find some 92 million "copy-org" occurrences -- in which there is not much of an inflationary factor due to organizational level and item identification problems. Extending this figure of 92 million by the study coverage factor, there are some 150 million copies of something received by somebody each year. That is not a comprehensible figure and is useful only to indicate the magnitude and complexity of the "community" system!

c. Interdepartmental Flow Patterns

Although small in proportion to internal flow volumes, the interagency flows are large in both absolute volumes and community concern. Figures III r (7-24) represent a picture of each agency in relation to its receipts from and then sendings to each other agency. The same biases and limitations of the other flow charts apply to these also. The accuracy of the statistics shown on the charts is no more than to one significant digit, much less the eighth digit as implied by the figures! There are two nontypical individual item flows that particularly distort three of the figures. In Figures III r (18) and (22), one item, STEP abstract cards which are sent by an Air Force government contractor (AID/LC) to multiple but unidentified USIB Departments, accounts for 45,000 "issues" and 4-1/2 million "copies" of the flow volume from "Air Force" to "other gov." In Figures III r (18) and (23), one item (translation "term" cards) provided by an Air Force Government contractor (AID/LC) to an Air Force nongovernment contractor (IBM) for a mechanical translation project, accounts for 432,000 "issues" and "copies" flowing from the Air Force to "non-Government." With these qualifications the following tentative generalized observations are made:

(1) CIA and State send out more than they receive in terms of "total copies."

(2) NSA is likewise so classed on the basis of "items" and probably would be on "issues" and "copies" if volumes of field produced "end products" were not absent from the data base. Even so, the amount of receipts by NSA is

TOP SECRET

TOP SECRET

25X1

very much the same order of magnitude as the other agencies in terms of both "issues" and "copies."

(3) Except for receipts by CIA, the total of receipts (both "issues" and "copies") is quite consistent by major agencies -- that is, about 1 million to 1.5 million issues and 3 million to 3.5 million copies per year.

(4) USIB interaction with "other" US Government departments looks appreciable in magnitude in the charts. As a group they would be like one more member agency. Including Commerce, they receive some 1.5 million "issues" and 3.5 million "copies" per year from USIB agencies. They send to USIB agencies some 2 million issues and 4 million copies. Also, further analysis is required to verify what portion of the "other US Gov't" elements are doing contract work for one or more of the USIB agencies.

(5) The "nongovernment" flow relationship (except the special translation cards) is almost negligible in comparison with the other charts. "Access" by rather than "dissemination" to contractors could explain this factor, which occurs primarily in the Air Force.

(6) As expected, the bulk of flow with allied foreign intelligence services is with CIA, NSA, and State.

wise the receipt and dissemination by the Clandestine Service with foreign governments is not in the data base.

(7) The flow from foreign governments to USIB agencies is somewhat higher than to foreign governments in terms of "issues" and "items."

(8) Figures r (7) and (8) indicate that CIA serves as an acquisition funnel for the community in that nearly half of its receipts are from non-USIB elements, whereas only about one-sixth of its distribution is to non-USIB agencies.

(9) Total "receipts" activity in Figures r (7-24) is about 10.4 million issues and 34.2 million copies per year. Of this activity, about one-eighth (1.6

TOP SECRET DINAR

TOP SECRET

million issues and 4.3 million copies) of the flow activity by USIB agencies is with non-USIB elements.

(10) It is assumed that most of what an agency sends to other agencies is its own product. The flow activity of CIA, for example, shows that some 17.5 million issues and 37.5 million "copy-org" occurrences of internal flow derive from some 5 million issues and 15 million copies, indicating about a threefold internal proliferation. Taking the aggregate of all six major USIB agencies on the same basis, we see some 30.6 million "issues" and 68.8 million copies of internal flow deriving from some 10.4 million "issues" and 34.2 million copies, similarly indicating a proliferation factor of only 2-1/2 to 3.

Recapping the over-all flow picture, we find great volumes moving between departments but four or five times this activity within departments; two departments send out appreciably more than they receive; external USIB flow is small in comparison with internal USIB flow but is appreciable in absolutes; and the over-all proliferation rate is low.

d. "World" Flow Patterns

Having wrestled with the big picture, it will be useful to view more specific sectors of the flow pattern in the community. It was always a possibility that division of the flow patterns by "worlds" would be useful for comprehension and analysis. Although it proved not possible within Stage I limitations to do this for all "worlds," this was done for two of the worlds: foreign publications and photography. In both cases as a result of the "world" treatment it was convenient to make integrated separate report sections out of them -- that is, orgs, people, items, flows, processes, problems, conclusions, and so forth. This separate integrated section on the "world" of foreign publications is contained in Appendix F. In the case of photography an added factor for separate treatment was security compartmentation. Thus the PHOTINT picture and discussion is contained in Appendix H, Volume VI. In addition to the two complete "world" treatments a large part of the flow pattern for the "world" of HUMINT, or Human Observation Intelligence, was obtained by analyzing the flow patterns of the "Information Report" type items in the various principal agencies. What follows here then is a picture of the flows of "IR's." The group of specific SCIPS items from which this composite flow picture was made consisted of

- (1) Department of State Airgrams
- (2) CIA CS Information Reports

TOP SECRET

TOP SECRET

25X1

- (3) CIA TDCS Information Reports
- (4) CIA OO Information Reports
- (5) AID Airgrams
- (6) NSA Information Reports (a very small unidentified group of collateral reports)
- (7) Air Force Information Reports
- (8) Army Information Reports
- (9) Navy Information Reports

Figure III s (1) shows that there are some 280,000 issues of these information reports (IR's) generated each year. (Not included in the charts are another 20,000 to 30,000 "IR's" per year produced by USIA, AID, [] and so forth.) The first thing that happens is that each agency reproduces its own IR's for internal and external distribution, as shown in Figure III s (2). This results in about 5 million copies for internal use and 7 million to 8 million copies for external use -- a total initial proliferation of about 45 copies per issue. Figure III s (3) shows which agencies are receiving this external flow in terms of "issues" and compares that volume with number of issues originated by each agency. This chart shows that everybody except State receives more IR's than it produces and by a ratio of about 3 to 1 over-all. The lowest producers have the highest ratios. The chart gives an impression of great multiplicity of receipt. Further analysis, however, shows that (not counting its own IR's) CIA receives 80 percent of all other agency IR's; that State receives 50 percent; that NSA receives 65 percent; that Army receives 50 percent; that Navy receives 60 percent; and that Air Force receives 50 percent, giving an over-all weighted average of 60 percent, which can be considered an over-all selectivity factor.

Figure III s (4) portrays, again in terms of "issues" not copies, Information Report interest by each agency including its own. The popularity of any one agency's IR's is indicated by the varying width of the color bands. In general, the selectivity factor seems to vary more by selecting agency than originating agency -- for example, NSA selects about the same proportion from each agency's production. Although Figures III s (3) and (4) indicate that either there is great dependence on each other's field collection or that there is great duplication of departmental interest and processing, the charts do not indicate which is the case or to what degree. There are other interesting observations to be made from Figure III s (4). For example, although CIA is the largest consumer of State IR's, State is next to the smallest consumer of CIA IR's, and the proportion of the total IR data base available to each agency, including its own, is as follows: Air Force, 50 percent; Army, 53 percent; Navy, 62 percent; NSA, 63 percent; State, 71 percent; and CIA, 89 percent.

TOP SECRET

25X1

TOP SECRET

25X1

Figure III s (5) portrays the same picture as Figure s (4) except in terms of "copies" rather than "issues." With the exception of CIA the flow pattern is about the same for "copies" as "issues," with the average number of copies per issue sent to each agency as follows: NSA, 5; Army, 5; State, 7; Air Force, 9; Navy, 10; and CIA, 18. Thus we have the first indicator of varying internal processing procedures or varying degrees of multiple utilization.

e. Selected Item Flow Patterns

In order to complete the flow picture, some of the items selected for analysis (see 2, g, above) were traced on through the second level of dissemination. It is reasonable to project that there is for most items a third and fourth level of further dissemination, but the scope of the study precluded in nearly all cases the survey of those activities. Figures III t (1-12) depict the flow patterns of specific items, selected for variety but not necessarily representative. Figures III t (1-4) are an extension of the HUMINT, or IR world picture. The internal dissemination within Army and Navy of Information Reports other than their own was not separately identifiable in the survey data. Although there are some obvious gaps and conflicts within the flow patterns, the charts do depict the breadth of dissemination as well as how many and which organizations are involved in the second-level dissemination process.

Figure III t (5) portrays the flow pattern for an item in the "initial intelligence" world. The total number of copies received per year by each organization is indicated. The colored geometric symbols represent processing of the item discussed in the next section -- the organizations where these appear are those surveyed during Stage I. As can be seen by the question marks, this is far from a complete picture. The varying number of copies received by apparently equivalent commands raises some questions about third and fourth level reproduction and dissemination. Figures III t (6) and (7) show the patterns for two selected items in the ANALINT, or finished intelligence world. Figure t (7) shows the average number of copies per issue received by each organization rather than total copies per year. This item, the CFTIB, is published in some 16,000 copies. Figure III t (8) shows what might be the typical distribution of a published REFINT, or reference-type item, the CIA IPI. This chart shows that SCIPS surveyed 11 of the some 90 organizational elements shown on the chart compared with 5 of the nearly 100 elements in Figure III t (7).

Figures III t (9-12) show the distribution pattern of fairly representative SCIPS items in the Communications Intelligence world. The difference in this world and in the PHOTINT world

TOP SECRET

25X1

Approved For Release 2007/11/20 : CIA-RDP80B01139A000600020001-5

Page Denied

Approved For Release 2007/11/20 : CIA-RDP80B01139A000600020001-5

TOP SECRET

25X1

No. 35. The "Army IR" is sent by 115 orgs (including attachés) and received by 67 orgs, each on the average receiving about half of the issues in an average of 34 copies resulting in some 3 million copies per year.

Recapping the flow picture derived on an individual item basis, we find multiple levels of dissemination, a great number of different organizations involved, and some items resulting in astronomical numbers of copies around the world.

4. Processes

a. General

Having "determined" which organizations are getting what and how many items, the only remaining piece of the picture is, "What do they do with them and how?" The restriction of the study effort for Stage I consciously excluded the study of the methods or techniques -- the "how" of processing. It was the intention and hope in Stage I to identify the processes (that is, name them) that were done to which items in which organizations. Although many attempts were made to pre-establish the identity of these processes for field survey purposes, no sufficiently common understanding of terminology between surveyors was achieved. The revised survey system, therefore, permitted free textual description of the processes found on site. During the post-survey editing phase these narratives were coded with some 30 process codes. (For further discussion, see Appendix G, Volume V, and for the process codes, see Exhibit n.)

b. Process Descriptions

The 318 processing activities are described in Appendix C, Volume III. All the disadvantages of textual reporting are evident in the full process descriptions -- however, they are believed to be informative as to the kinds of processing being done in the community, even though the accuracy, completeness, and intelligibility of any one description may leave much to be desired and often raises more questions than it answers.

c. Over-All Picture

The standardized process codes as applied by the editor were used to correlate items receiving common processing. Within the accuracy of the surveyors' narrative description this correlation is made in SCIPS unpublished Data Catalog P-1. Unfortunately there has

TOP SECRET

25X1

TOP SECRET

25X1

been no over-all analysis of this portion of the data due to time restrictions. The computer file for this "processes" data comprises some 78,000 records and, when printed out, comprises some 6 volumes. 37/ An analysis of over-all processing should show the quantity of items receiving each of the "processes" -- for example, how many different items are extracted? The mechanized data system and item identification problems precluded this kind of lengthy analysis during Stage I. The first indication of over-all processing is reflected in previous sections -- for example, by the flow picture: every org-copy occurrence implies some processing. The only other over-all picture of processing developed during Stage I is contained in Figures III u (1) and (2), where the occurrence of the processes in the various 42 organizations and some 300 processing activities are depicted. Some of these processes are not too distinguishable, such as "distribute" versus "disseminate," "abstract" versus "extract" or "summarize," and "analyze" versus "evaluate." Other processes are believed to be more common than the charts indicate, such as "destroy," "write," "select," and "sort" -- because they are so common that they were not noted consistently during the survey. Some of the processes are of special interest as follows:

- (1) The commonness of filing -- 98 percent of the orgs and 60 percent of the "activities" do filing.
- (2) The confirmation of the previous "equipment inventory" indication that reproduction and typing is widespread.
- (3) The number of activities that abstract/extract/summarize, perhaps 50 different activities.
- (4) The number of different organizations and activities that "punch" (20 and 50).
- (5) The low ratio of activities to organizations that do indexing, meaning at least some centralization of indexing within "organizations."

d. Selected Items

Although an over-all analysis of the processes could not be made, some work was done for some individual items. Figure III u (3) shows the number of different organizations in which some selected processes occur on six of the Information Report items discussed in

TOP SECRET

25X1

TOP SECRET

25X1

section 3, d, above. The pattern here is that each of these items receive almost all of these processes by about six different organizations except that many more organizations than that do "dissemination" and "filing" of IR's.

The processing of selected specific items also is shown in Figures III t (5-8). The extensive filing of items is reflected in the case of the daily Current Intelligence Digest and the Intelligence Publications Index, but generally the processing reflected on these charts is not considered adequate or sufficiently accurate for analysis.

e. Selected Processes

Two particular "processes" are keys to community IP: indexing and filing. These two processes were studied in greater depth than the other processes. The "indexing" picture is reserved for discussion in the "problem" section of this report. The filing processes warrant special treatment and receive it in the following sections.

f. Selected Organizations

Although internal processing methods were not studied in Stage I, the brief process descriptions enable the drawing of rough process charts. The processing in four organizations are so depicted in Figures III u (4-9). In FTD as shown in Figures III u (7-9) are the three typical degrees of information control: indexed documents; indexed abstracts and extracts; and extracted formatted event information, respectively.

5. Files

a. General

Filing is one of the two processes that were studied in more detail than the other processes. Filing can also be considered a part of the flow pattern -- that is, temporary storage. In addition, filing is the other part of the picture of items and flows: "Where can it be found?" or "Who keeps what?" Permeating this files picture are all the key questions and clichés of "unique data base," "data exchange," "standard format," "automatic input," "storage and retrieval," "common indexing," "all-source index," "man-machine systems," "natural language processing," and so forth. Because of its centrality, considerable study effort was placed on files. There are many known variables connected with files, and all are of potential significance. What follows, then, is a picture of the files, as found in the organizations surveyed in terms of summary statistics

TOP SECRET

25X1

TOP SECRET

25X1

on occurrences of the variables and then (in subsections r, s, and t, below) correlations between those variables. The number of variables is large, and the potential correlations are multitudinous. Although all the correlations could not be made in the time available, it is believed that the most important ones were made and the result is a series of charts that in combination comprise the picture of community files (and, therefore, inherent processing and problems) within the scope of Stage I. Except for 116 files covered in Appendix H, Volume VI, all files surveyed are considered in this section.

Throughout the section, there are two bases for measurement and comparison: the number of "files" and the number of "unit records." A file is a matter of definition, as discussed in the survey instructions (see Exhibit g, Appendix G, Volume V). Although no single criteria was used to determine a file, one of the more important facets of file definition used in the survey was that of file "order" -- that is, any change in filing order results in a separate file, even though the file content is the same. Files of concern include index and process control files as well as substantive item files. A unit record is an item in file -- for example, a document, a card, a record on magnetic tape, a photograph, and so forth. Thus either or both the number of files and the number of unit records can be important in measuring a given variable or descriptive parameter of files.

There were about 20 major filing parameters that were measured during survey. The number of instances of "not recorded" during the survey usually were very few and varied slightly by file parameter. No effort has been made to adjust these variances -- thus the total number of files or unit records differs somewhat between charts. The principal concern is comparative proportions and orders of magnitude.

b. Total Files and Size of Files

For the 43 organizations surveyed (excluding the 13 organizations covered in Appendix H, Volume VI), 926 files were identified. The numerical distribution of these files by agency surveyed is shown in Figure III v (1). Figure III v (2) presents the variance between these agencies in average size of files, and that accounts for the difference between agency proportions in Figure III v (1) versus Figure III v (3), which shows some 220 million total "unit records" by agency. This makes the over-all average file size 240,000 unit records.

Regardless of how the files are categorized, there is no denying that 220 million is a lot of unit records to have in the

TOP SECRET

25X1

TOP SECRET

25X1

files, particularly when the analyst, archival, and remote location files are not included. This in itself gives an order of magnitude to the community IP problems. The major categories of community files that were generally excluded from the Stage I study should be borne in mind: (a) administrative (personnel, logistics, security, and so forth); (b) the multitudinous analyst files; (c) collection management files; (d) books; (e) archives; (f) counterintelligence, and so forth. The 220 million file items represent the major "central reference" type of files primarily.

The proportions between agencies in Figure III v (3) do not particularly reflect study coverage except the CIA preponderance. In the case of CIA, about half of the files and 80 percent of the unit records were in the Office of Central Reference -- which is to say that not all the studied CIA files are in OCR but that the big ones are. This should not obscure the fact that some 10 million records are outside OCR. Also, the big Walnut and RI file volumes were not provided by CIA. No particular significance is put on the average variance of the size of files between agencies, because of file definition, except to note that the extreme in State (and possibly Navy) does indicate a basic difference in filing philosophy -- that is, less segmentation. It is of interest to note that even the average size of files (some 240,000 records) is very large to the "outside world," where all the publicized documentation techniques are being developed and applied. Furthermore, it is suspected that there are not very many "average size" files -- that is, most files are either quite small (5,000 to 50,000) or huge, like 5 million unit records.

It should be realized that the foregoing over-all figures have added apples and oranges. For example, a file in the data base may range from 2 reels to 10 million punched cards. The sample is of such a size that proportions between specific variables as discussed in the following sections are credible.

c. Age of Files

Figure III v (4) depicts the age distribution of some 870 of the files (and the corresponding 150 million unit records) for which an effective date -- that is, year of establishment -- was recorded during the survey. This leaves a lot of files and unit records for which the date of establishment was not determined. These most likely would be added to the pre-1960 age groups. The 1963 group is not representative, because most of the survey was done in 1962, and the 1962 figures likewise are somewhat deflated. Effective file date does not directly measure the age of items in the files but does reveal some interesting trends in community filing. For example, one-fourth of the files were established within the last 3 years. The

TOP SECRET

25X1

TOP SECRET

25X1

oldest files (pre-1950) are the largest (500,000 unit records per file). Surprisingly the files established in 1960 are already of larger average size than those established in the 1955-59 period. Files established in 1961 are already nearly as large. This implies a trend toward larger files or that many files which were established in the late 1950's have not been growing much. It is probable that the trend is toward larger files. It also looks as if there is a growing trend in the number of files being established per year: 1950-54, 40 per year; 1955-59, 50 per year; 1960, 60 per year; 1962, 60 per year, with a bulge in 1961 to 75 files per year.

d. Present Rates of Growth of Files

With the present size and age distribution of files in mind, the rate of growth of existing files should be considered. The previous section and Figure III v (4) gives an indication of past rates of growth and trends. Figure III v (11) shows the present rates of growth of 606 existing files on which information on the rate of growth was obtained during the survey, including "zero growth" files. Aside from indicating that operators do not know and would not readily guess rates of growth for one-third of the files, no speculation is made on the nature of the other files as regards present rates of growth. Thus the nearly 20 million records per year at the present rate of growth must be considered minimum. The rate is some 30 million unit records per year when extended to 900 files. Again, the limited scope of study should be considered. The rough implication is a doubling of present holdings in 7 years, as a minimum. The variances in present rates of growth between agencies, as shown in Figure III v (11), are attributable more to the proportion of files considered than any other factor. The chief exception to this is the Air Force, where an exceptionally high rate of growth is indicated.

e. Future Rates of Growth of Files

In addition to estimates of present rates of growth of files, projected rates of growth for existing files was sought during the survey. The information was obtained for only 395 of the existing files, including "zero" growth and even some "negative" growth -- that is, a decreasing file. Again, it is not known what speculation to make about the future rate of growth of the other 500 to 600 files other than that no one knows or can readily estimate future rates of growth of these files. Figure III v (12) shows the projected annual rates of growth in terms of unit records by agency. The picture is quite similar to "present rates of growth" except that the average per file is about 38,000 unit records per year versus about 32,000 unit records per year for present rates of growth.

TOP SECRET

25X1

TOP SECRET

f. Purge Criteria

In the context of the size, age, and rates of growth of present files, an obvious question is, "What is the file purging practice in the community?" Five purge criteria were established prior to the survey. Information on the applicability of these criteria was obtained on a good sample: 871 files containing nearly 200 million unit records. Figure III v (13) shows the proportionate distribution against these criteria for all departments. The distinction between three of the criteria -- "supersession," "correction," and "evaluation" -- is considered questionable, and so they are combined in this summary presentation. The striking features of the file purging picture are (1) the majority of files are not purged, and (2) of those that are purged, very few are on a direct file activity basis. The pattern by agency is shown in Figures III v (14-19). The most striking variances by agency are (1) the high "supersession" proportion in the Army and (2) the high "date" proportion in the Navy.

g. Files with Standard Formats

One of the important distinctions to be made in considering the community's files is the question of whether the items or unit records in a given file are of uniform format or whether the file contains items having a variety of formats. In addition to standard formats being one of the very few specific areas of IP where USIB has pronounced a policy, 5/ it bears directly on the assigned problem areas of "indexing" and "data exchange." For this reason, additional survey facts were obtained to describe formats in those files that contained only one (or a few) formats for the unit records. This unit record format information is presented and discussed in the "problem" section of this report. Here in this section for "picture" purposes there is Figure III v (20), which presents an over-all count of files and unit records with and without standard unit record formats. This chart shows that 36 percent of the files and 42 percent of the unit records are of standard format -- that is, only one or a very few different formats in the same file. State and Navy had few files and unit records in the standard format category. The other departments ran higher than the 36 to 42 percent. On an over-all basis the average size of the files was only slightly higher for standard format files than for nonstandard files. In addition to the 364 files tallied, there were 12 other files surveyed with standard formats, but the file size -- that is, the number of unit records -- was not obtained. These were relatively large files (Walnut, Finder, and RI Main index). As a generalization, it can be said that nearly half of the files consisting of some 100 million file items do have internally consistent formats as called for in the USIB policy.

TOP SECRET

TOP SECRET

25X1

h. Manual Versus Machine Files

USIB has a policy calling for mechanization of reference files. 5/ For 926 files the method of file maintenance -- that is, manual or machine -- was noted during the Stage I survey. "Machine maintained" meant not only computer and punched card files but also aperture and microfilm reel files but did not include a mechanized elevator file. Thus the portion of all files and unit records labeled as machine files in Figures III v (21) and (22) overstates mechanization in terms of the proportion of file information that is machine-readable. The charts show that about 30 percent of the files containing 37 percent of the unit records are machine-maintained. Even though some of the large machine files (Finder, Walnut) are not included, some 926 files and 220 million unit records are in the sample. About half of the 78 million unit records in machine-maintained files are punched card index files to hard copy document files. Bearing in mind the limited study scope, these are magnitudes that one must consider when contemplating an "all-information" computer file or a single "all-source" index. The proportion variances between agencies in Figures III v (21) and (22) somehow do not seriously disturb previous intuitive estimates, but likewise the magnitudes were not previously appreciated.

i. Types of Files

During the survey the "type" of files was recorded as follows:

Master: that is, the original file for the items.

Derivative: a file created from another file but having either only a part of the "Master" file or the whole of the "Master" file but in a different order (usually to provide alternate means of entry to the same information).

Duplicate: a complete copy both in content and in order of a "Master" file (usually done to enable multiple servicing or backup).

Suspense: a file of items yet to be merged with the master file.

Figures III v (23) and (24) depict the proportions by type and by agency between these types for 926 files.

TOP SECRET

25X1

TOP SECRET

25X1

It is rather surprising to find so few derivative and duplicate files. If true, then either the utilization rates must not require multiple access (because not very many of the files are computerized to permit multiple access to a single file) or a new file is produced rather than derived. Backup duplicate files may be in archives where the survey was not done. It is suspected that there are many more "suspense" files than indicated but that they are not usually considered "files." Again, proportions are not the whole story -- for example, the absolute figure of 40 million unit records in derivative files is appreciable. Why the size of derivative files ran twice as large as master files (400,000 versus 200,000) is answerable only by identifying and analyzing the individual files. Variances by type between agencies is not particularly striking except perhaps the absence of derivative files in State.

j. File Content Categories

A complete catalog of the files in terms of substantive content would require an inventory of all 220 million items in files. This is patently impossible. As a substitute, SCIPS recorded on the Process Description Sheet what "items" are currently going into each file. Exploiting this information requires more analytical time than was available in Stage I. Some insight into file content is revealed by the file title (see Appendix D, Volume III) and by the narrative description of filing criteria 36/ recorded during the survey. Aside from substantive content, files need to be categorized as to whether the items in file are documents, abstracts, information, index information, or combinations thereof. In Figures III v (25) and (26), 926 files and 210 million unit records have been plotted by the category most representative of items in files. Items representing a combination of categories were assigned to the "mostly" category (there were 1,165 categories tallied for 926 files). These charts show that there are more document files than any other category but that there are more index unit records than any other category, and more "Abstract/Extract/Summary" type items than documents. One surprise is to find some 170 "data" files with 21 million records. "Data" file here is generally intended to mean extracted and formatted information rather than documents or pieces of documents or reference to or description of document content. The ^{virtual}absence of CIA in the "data" files category may be partially but not wholly explainable by inconsistent discrimination and definition during the survey. Index files are more than twice as large as document files, and data files are small in terms of numbers of unit records (125,000 per file), perhaps because they are newer files.

TOP SECRET

25X1

TOP SECRET

25X1

k. Purpose of Files

The categorization of files by purpose served should be a useful management tool. "Purpose" here means, in terms of functions to be served, the reason a given file was established or maintained. Seventeen such purposes were established for the files survey [see Exhibit g (4), Appendix G, Volume V]. The occurrence distribution of these "purposes" is shown in Figure III v (27), with the 17 purposes combined into 6 groupings. The high proportion of "regular service" files versus "special project" files means that system-oriented files (as opposed to information-problem-oriented files) are predominant. Although not shown in Figure III v (27), the survey data did reflect some variances between agencies. Only one-tenth of Navy's files were "regular service"; nearly 95 percent of State's files were "regular service." NSA had the most "special project" files. Variances in the surveyors' definitions of terms should be checked out before drawing further conclusions.

l. Intended Utilization of Files

During the survey it was determined for some 900 files whether the files were established and maintained with the intention of serving (1) any and all USIB agencies, (2) only components of the holding agency, or (3) only the local component. The resulting survey data is shown by the number of files in Figure III v (28) and by the number of unit records in those files in Figure III v (29). Although restricted "intent" is not necessarily restricted "access" to the same degree, it is some indication of this also. Likewise, broad "intended utilization" does not necessarily imply a commensurate responsibility to provide service but again implies some degree of formal or informal sense of responsibility. Neither does "intended" reflect the degree of "actual" utilization. With these definitions of "intended utilization" the response bias probably favors the "USIB" category. In any event, even in the "central reference" type of organizations surveyed, only about one-half of the files are intended for utilization for any and all USIB agencies. Well over one-half of these files and three-fourths of these unit records are in CIA, most of the rest being in NSA with a little in the military agencies (presumably those files of their own products). Another point of interest, considering the definition bias, is that more than 200 of the files and some 20 million unit records, or more than 20 percent, are intended for use only by the local component.

m. Security Classification of Files

There are two aspects of security classification with respect to files: (1) the security classification of the over-all file and (2) the various security classifications of the items in the file.

- 36 -

TOP SECRET

25X1

TOP SECRET

25X1

Both factors were recorded during the survey. Figure III v (30) shows the distribution of files by four groupings of over-all file classification. Addition of the files covered in Appendix H, Volume VI, would just about double the fourth group. The over-all file classification is almost an absolute access determiner. Although it looks small on the chart, there are well over 100 unclassified files. The most noteworthy feature of Figure III v (30) is the absolute number of files and unit records with restricted access (nearly 500 files and 60 million unit records). The question is, Are all the items of information in those files likewise classified, and how unique are they? Figures III v (31) and (32) show the files and unit records by classification of items in files. The mixed classification groups represent files where access to some items is more restricted than the classification of the items requires because they are co-filed with higher classified items. The first three single-classification groups in Figures III v (31) and (32) amount to some 500 files containing some 140 million unit records. In the "CST" grouping the amount of TOP SECRET files is negligible. The preponderance of mixed classified and unclassified files suggests that no one is particularly concerned about servicing on an unclassified basis. The proportions within groups is not known. Noteworthy variances between agencies are as follows: (1) the majority of NSA files being all classifications (except p), (2) the almost complete separation of special intelligence items in CIA, (3) the exclusive one-category of files for State. The extent of mixing classifications would seem to indicate that security classification is more of an access problem than a filing criterion.

n. Dissemination Controls on Files

In addition to the over-all security classification of files, the survey recorded any dissemination controls (DCID 1/7) on the files. The response is indicated in Figure III v (33). The number of different controls (both old and new) itself poses a file control problem of some magnitude, but, like security classification, it probably is not a filing criterion -- that is, separate files are not established for each dissemination control. The two largest groups in Figure III v (33), "All Dissem Controls" and "Unspecified (classified files)," accounting for more than one-half of the files really means that in terms of dissemination controls file content is unknown and probably uncontrolled in terms of indexing or formatting. The presence of specific controls such as "No Foreign" and "No Contractors" is significant, but their absence is as much a matter of "undetermined" as anything. These two controls plus "own agency only" are the only significant specific controls evidenced in Figure III v (33). Variances between agencies are believed to be a result of survey factors and non-pre-determination of dissemination controls on items.

TOP SECRET

25X1

TOP SECRET

25X1

o. Consumer Access to Files

The surveyors recorded for some 850 of the files (192 million unit records) whether consumers served by the file had direct physical access to the files or whether they had access to the file items only through a file servicer or an intermediary. As indicated in Figures III v (34) and (35), only 30 percent of the files and only 15 percent of the unit records are directly accessible to the consumer. For 11 percent of the files and 3 percent of the unit records, there is no consumer usage, meaning they are intermediate or process control files for the use of system operators. There were three striking variances between agencies: (1) the high proportion of "direct" access files in NSA, (2) no "direct" access files in the Navy, and (3) the relatively high proportion of "no access" files in the Army. The study of more units in the Navy (such as STIC) would reveal some direct access files. "Direct access" files are less than half the size of "indirect" files on the average, and the "no access" files are relatively quite small. The preponderance of "indirect" access files is not too surprising considering that the sample consists primarily of "central reference" files rather than the thousands of "analyst" files. "Indirect" access does not in itself guarantee file integrity; rather more probably it is a reflection of necessary search assistance.

p. Physical Form of File Items

During the survey, each file was described in terms of file housing, with the unit record being the first order form. For example, in a given file the first order of file housing, the unit record, might be a "5 x 8" card, the second order a tray or drawer, and the third order housing a safe. Although no analysis has yet been made of second and third order housing, the files have been analyzed in terms of the form of unit records. During analysis the great variety of unit records was normalized into 14 groups now called "file forms." The proportions of files and unit records surveyed between these different "forms" are shown in Figures III v (36) and (37). Two CIA aperture card files totaling about 3 million unit records were recorded as punched card files. Corrections for this would not affect the punched card category very much but would raise the ranking of aperture cards on Figure III v (37) to sixth or seventh place. It would be interesting to have a time correlation over the past 10 to 15 years for the "big five" forms: punched cards, 5 x 8 cards, documents, 3 x 5 cards, and magnetic tape. "Document files" can be termed "hard copy files." The popularity of 3 x 5 card files is surprising in that library card catalogs generally were not included in the survey, nor were the large book libraries. All the "big five" forms except "documents" imply quite a lot of transformation

TOP SECRET

25X1

TOP SECRET

25X1

of information items, as do the "minicard" and "4 x 6 card" forms. Punched cards, 3 x 5 cards, 5 x 8 cards, and magnetic tape comprise some 158 million records, which is some 75 to 80 percent of all file items in the sample. Only punched cards, magnetic tape, minicards, and listings imply machine-readable information, and this is only about 80 million records. Of the big five forms, probably none is decreasing in use unless it is hard copy documents. Presumably, punched cards are at their peak in popularity, and magnetic tape is coming to the fore. The effect of file definition on the "average file size" by form shown in Figure III v (38) is unknown. The average file sizes for the first nine forms are all so large that it is hard to consider the variances. It should be noted as a generalization that "hard copy" files are smaller in average size than all other forms of transformed or mechanized files except microfilm reels of documents. In analyzing the minimum and maximum file size variances by file form, it was found that averages give the wrong impression. For example, although the average file size for 3 x 5, 5 x 8, 4 x 6, and punched card files is around 300,000 unit records per file, there are files in all these forms of 2 million, 3 million, 4 million, and 5 million records in each; and in the case of punched card files there are up to 10 million records in one file. Magnetic tape files run up to nearly 2 million records. The variances in popularity of file form by agency are shown in Figures III v (39) and (40). The more notable points are as follows: (1) CIA dominance in punched card files, (2) the bulk of "documents" in State and CIA, and (3) Air Force and NSA dominance in magnetic tape records. In terms of future automation as called for in the USIB policy statement, 35/ the notable figure is the number of reference files in the form of non-machine-readable cards (3 x 5, 4 x 6, 5 x 8): more than 80 million.

q. Order of Files

For survey purposes in describing the sequence of records in files, 10 criteria for file order were established: subject, area, persons, date, source, title, locally assigned accession number, item serial number, security classification, and dissemination controls. The survey form [see Exhibit g (4), Appendix G, Volume V] was designed to record up to five filing orders for any one file. An example of five filing orders would be documents filed, first, by geographic area (USSR); second, by subject category (Politics); third, by individual person's name ; fourth, by date of information (1960); and, fifth, by document title. The occurrence in the data base of files by first filing order is shown in Figures III v (41) and (42). The difference in relative rank between these two figures is explained by variances in average file size shown in Figure III v (43). The outstanding features of these three charts are: (1) the high occurrence of subject, area, persons, and source as criteria for

25X1

TOP SECRET

25X1

TOP SECRET

25X1

first filing order; (2) the relatively low occurrence of files with date or title as the first filing criterion (even so these amount to some 88 files and 7 million unit records); and (3) the large average file size for "persons" and "source" files. Filing orders included in the "other" category are usually subelements or combinations of the other categories but can be such things as file "form." "Serial numbers" often include source and date information; "accession numbers" were not always distinguished from "serial numbers" during survey. The frequency of the category "area" may include an indication of organizational structure in the organization being surveyed more than a file order as such.

Variance in filing order between agencies is reflected in Figures III v (44) and (45). According to these charts, nearly all agencies use all the criteria for the first order to some degree, and there is more similarity than dissimilarity in proportions between agencies. NSA and CIA seem to favor one category over another less than the other agencies. State is heavy on subject files, and the Air Force seems to favor area files.

Although all successive orders have not been analyzed yet, the second order was examined in the two cases where "area" and "subject" were the first file order. For the nearly 300 files where the first filing order was "area" about 30 percent of the files and 20 percent of the unit records had no further order -- which is to say they were "pure" area files. However, for the 189 "subject" files nearly 30 percent of the files and 50 percent of the unit records were not further ordered. Of the some 200 "area" files that had a second order, 106 of them containing 27 million unit records were ordered second on subject. If the first order "area" occurrence reflects organization (for example, the "German Section") then subject is more frequent as a file order than area. In any event, they seem to be about equal in usage as first filing order. The other noteworthy result of analyzing the "subject" and "area" ordered files is that "persons" is second order for 49 "subject" files (18 million records) and for only 9 "area" files (0.7 million records). Adding these to the first order "persons" files, we find a large number of files (118) and a very large number of records (40 million) ordered on "persons." These 40 million records may be considered as constituting the community's biographic files (excluding counterintelligence and security). Biographic files are explored in further detail in the problem section following.

It is in the second order of subject files that the expected large number of files ordered by date are found (either document date, information date, or receipt date) -- 42 files containing 30 million unit records. Within the first order area files we find 14

- 40 -

TOP SECRET

25X1

~~TOP SECRET~~ []

files containing 7 million records then ordered by serial number. Accession number does not seem to occur much as a criterion except when it is the first order.

r. Correlation of Machine Files with:

There is, of course, a direct correlation between the physical form and whether or not a file is machine-maintained or manually maintained. Of the 14 file form categories the following are most usually machine-maintained: punched cards, aperture cards, microfilm reels, minicards, and magnetic tape. Figures III v (21) and (22) discussed above show that about one-fourth of the files and one-third of the unit records are machine-maintained.

(1) Consumer Access

Figures III v (46) and (47) show for each method of consumer access ("direct," "indirect," and "none") what portions of the files are machine versus manually maintained. As can be seen, very few of the machine files (8 files and 250,000 unit records) have direct consumer access.

(2) Types of Files

Correlation of "type" of file (master, derivative, duplicate, suspense) with manual versus machine-maintained shows that about half of the "derivative" files are machine-maintained versus one-third of all files being machine-maintained.

(3) Content of Files

Figure III v (48) shows the correlation of file content versus method of file maintenance. The proportion of index and data files that are machine-maintained is even higher than expected. At the same time, the lack of mechanization of abstract/extract files is noteworthy. Even though the majority of index records are machine-maintained, there are still some 35 million index records manually maintained. Figure III v (49) shows that the only machine-maintained abstract/extract files are in CIA, as is the bulk of the machine-maintained (mostly punched cards) index files while the Air Force and NSA hold the bulk of the machined data files.

s. Correlation of the Rates of Growth of Files with:

During analysis, both the present and the future rates of growth of files were correlated with security classification and

~~TOP SECRET~~ []

TOP SECRET

25X1

"Form" of unit record. Correlations of rates of growth with file order and content categories might prove beneficial also but have not yet been made.

(1) Security Classification

Figures III v (50) and (51) compare the present and projected rates of growth of files in each agency by over-all security classification of the files. (Files in Appendix H, Volume VI, are not included.) Although the sample size is different for the two factors, both present and projected rates of growth are proportionate to present size regardless of classification, except that "Special Intelligence" files have a proportionately higher rate of growth than CONFIDENTIAL/SECRET/TOP SECRET files. Also for Special Intelligence files the projected rate of growth was the same as the present rate of growth even though the sample is 200 files smaller; so the trend must be toward files with over-all classification of "Special Intelligence." This is particularly noticeable in CIA and the Air Force. The operational areas in the Army, the Navy, and State where these files occur were not covered during the survey. Although the trend is one way, there will evidently always be large numbers of files and records in the CONFIDENTIAL and SECRET classification (8 million to 10 million more such records every year).

(2) Security Classification Versus Form

Figures III v (52) and (54) show the correlation of present rates of growth of files for file form and file classification. Figure III v (52) indicates that 5 x 8 cards are a favorite form for unclassified items, probably mostly from foreign publications. Likewise in the CONFIDENTIAL and SECRET files the 5 x 8 card form is prominent. In the Special Intelligence files [see Figure III v (54)] magnetic tape is a new factor. In all three security groups the popularity of 5 x 8 card files is striking and is some indicator of need for context in information files if complete data files can not be had. That is, a 3 x 5 card will not do. (This is a generalization to be sure.) The absence of microfilm reel files in the special classification is noted. Projected rates of growth in the various classification groups as shown in Figures III v (55) and (56) have much the same pattern as present rates of growth. The sample for projected rates is only half that for present rates and is insufficient for unclassified files to be included here. The projected annual rate of growth for CONFIDENTIAL and SECRET punched cards is about 10 percent. The rate of growth for magnetic tape records in Special Intelligence files is 28 percent per year. The projected rate of growth for 5 x 8 cards is higher in Special Intelligence files than in CONFIDENTIAL and SECRET files.

- 42 -

TOP SECRET

25X1

TOP SECRET

25X1

t. Correlation of the Form of Files with Other Factors

The file form (for example, punched cards, hard copy documents, and so forth) is the first factor in indicating the nature of processing techniques and often the first consideration in data exchange. How this factor, file "form," correlates with other filing parameters is discussed below.

(1) Security Classification

Figures III v (57-60) present the picture of surveyed files by each form within each of the groupings by over-all security classification. The popularity of the 5 x 8 card form over punched cards or 3 x 5 cards is reflected in Figure III v (57) for unclassified files. Book library card catalogs generally were not in the data base. The 1 million unclassified magnetic tape records is noteworthy. The proportions between file forms for CONFIDENTIAL and SECRET files (there are very few TOP SECRET files) shown in Figure v (58) set the physical form pattern for all files because CONFIDENTIAL/SECRET files comprise the vast majority of files. In this classification grouping, all forms occur including "strips" and "discs" in the "others" group. The big differences in the "Special Intelligence" files forms are the absence of 4 x 6 cards and photographs, the reduced proportion of hard copy "documents" and "micro-reel" forms, and the increased proportion of files and records in "magnetic tape" form. The sample for the compartmented files is small and not necessarily representative; however, the combined use of many different forms is noteworthy.

(2) Effective Date

Figures III v (61-67) attempt to show whether there is a general trend over time in the popularity of the various physical forms for file items. These charts are based on the effective date of establishment of the files. Therefore, the trend over time of the number of unit records in the various forms may represent not a trend of desirability but a limitation of prior commitment that cost of conversion cannot overcome. The number of files established in a given time period is more a reflection of the "any form desired" trend. In all charts the figures for 1962 and particularly for 1963 are incomplete. Figure III v (61) indicates a continued strong upward trend for punched card files since 1955. Figure III v (62) shows a decided downward trend for 3 x 5 card files (again excluding library card catalogs) to where virtually no new 3 x 5 files are being established. Figure III v (63) indicates that, except for three small files in 1961 and one in 1962, the 4 x 6 card form is even more passé than the 3 x 5 card form. Figure III v (64) shows that although the greatest popularity for the 5 x 8 card form was in the 1950's, and even before 1950, new files in this form continued to be established. Although the recently established 5 x 8 files look small in terms of unit records, they are comparable

- 43 -

TOP SECRET

25X1

TOP SECRET

25X1

in size for their age with the huge old 5 x 8 files. Figure III v (65) indicates that a lot of new hard copy document files were established in 1960 and 1961. Otherwise, there seems to be a "steady" trend in establishing new hard copy files. The more significant factor for hard copy files seems to be that there is sort of an upper average size limit (since 1950) of about 100,000 records. Presumably these files are closed and new ones established, whereas for the other file forms they just keep adding to the file. This indicates that the hard copy document files serve a different function than other forms -- which in turn imply information extraction to varying degrees. Although a strong trend in the use of document microfilm reels as a file form was expected, Figure III v (66) does not so indicate. Perhaps complete data for 1962 and 1963 would confirm the trend suggested by the 1961 figures of four files and 720,000 records. The older magnetic tape files indicated in Figure III v (67) probably represent file conversions. Except for incomplete 1963 figures, there is a strong upward trend in the use of magnetic tape, although not quite so strong as expected. Files in the "unknown" effective date grouping probably should be allocated to recent years. This would strengthen the indicated trend. The other file forms were of too low occurrence to discern meaningful time trends, except possibly maps and charts, which showed some increase in recent usage. In summary, it appears that all forms are "down" except punched cards and magnetic tape, with a strong upward trend, and 5 x 8 card and hard copy document files, which continue relatively steady. Although a thorough analysis of form-trend variances between agencies was not made, it was discerned that they generally follow the form preferences reflected in present file holdings shown in Figures III v (39) and (40).

(3) Purge Criteria

The over-all purge criteria picture was shown in Figure III v (13) and discussed in paragraph f, above. Although not charted, some analysis was made of variances in purge criteria for files by the various physical forms of files. Punched card files are purged more than the average for all forms, and there is a high incidence of purging on the basis of supersession/correction/evaluation as well as by date. Surprisingly, there is practically no purging of punched card files on an "activity" basis. What little purging there is on an activity basis occurs mostly in hard copy document files. Likewise "activity" was not a criterion for purging of magnetic tape files. What little incidence there was of purging magnetic tape records (only 15 percent of the magnetic tape files) was on the basis of "supersession/correction/evaluation." The higher "no purge" occurrences were in the magnetic tape and 3 x 5 card files. Purging on the basis of supersession/correction/evaluation occurs most frequently in 5 x 8 and punched card files. Variances in purge criteria by file form between agencies were not analyzed. In summary, two observations can

- 44 -

TOP SECRET

25X1

TOP SECRET

25X1

be made: (a) once information has been extracted or put under more than documentary control, the tendency is to never purge, and (b) usage rates are not a purge criterion in mechanized files where such bookkeeping and file reorganization capability is supposedly an advantage.

(4) Order of Files

What influence the physical form has on the order of files, or, vice versa, to what extent the order of files dictates physical form, is an important correlation for purposes of long-range planning and future systems design. Figure III v (68) correlates, in terms of the number of unit records, the principal file forms with the major categories by which files are first ordered. Thus it appears that punched cards are the favorite form for subject-ordered files, but Figure III v (69) shows that more punched card files are ordered on area than on subject, and Figure III v (70) shows that more 3 x 5 card files and unit records are in subject order than in any other one order. Figures III v (70) and (71) as well as Figure III v (68) show that whereas 3 x 5 card files are ordered by subject, the 5 x 8 card files tend to be ordered by area. Although that correlation is not too strong, it is of importance in view of the trend in usage of these two forms discussed in (2), above, indicating that 3 x 5 card files are dying out and that 5 x 8 card files continue to be established. Other correlations of interest are (a) most of the unit records ordered by date are magnetic tape records, and magnetic tape records [see Figure III v (73)] are usually in area, document number, or date order; (b) although the great bulk of hard copy documents seem to be filed by subject or area, Figure III v (72) shows that there are a great number of relatively small document files ordered by source and most every other order; (c) although there are more 3 x 5 cards filed in title order, there are 20 document files (versus seven 3 x 5 card files) that are ordered on title. Examination of the second filing order within the "subject files" [see Figure III v (74)] shows no change in form correlation except that many of the 5 x 8 and hard copy document files are ordered by date within subject. An examination of second filing order within area files [see Figure III v (75)] strengthens the correlation between punched cards and magnetic tape files with area-ordered files. Also, the 5 x 8 card form correlates more with subject-ordered files than with area-ordered files. A look at the files that have "persons" as either first or second filing order, as in Figure III v (76), confirms the proportions between file forms -- that is, all forms are used extensively for personality files with the bulk (58 million) being on 5 x 8 cards.

Although no charts are included, the other less frequently occurring forms were analyzed for correlation with file order. Aperture cards usually are filed in accession number order; machine listings usually are not in area order; maps, charts, photographs,

TOP SECRET

25X1

TOP SECRET

25X1

minicards, and folders usually are in area or accession serial number order; 4 x 6 card files are in area or persons order; and microfilm reels are in subject or accession number order. Although little analysis was made of interagency variations in form-order correlations, it is suspected that there are even stronger correlations within than between departments. In summary, there seems to be a significant relationship between file form and file order. That is, machine-readable records and composite forms usually are ordered on area, while non-machine-readable but extracted information records usually are ordered on subject, with all forms being used for personality files and hard copy document files being filed in a variety of orders. Viewed another way, when the physical form provides only one access point, the choice is "subject," and when the form provides multiple access (punched cards, magnetic tape), the normal order more often is by area. As a generalization, there are more subject categories than area categories, so that in very large files the subject has to be used to get sufficient initial discrimination.

(5) Content of Files

Figure III v (77) shows a clear correlation between physical form groupings and content categories. Again, in Figure III v (77) files are double-counted where file items are a combination of categories (about 1.3 categories per file). Index files are mostly punched cards, 3 x 5 cards, or magnetic tape (in that order), and most punched cards and 3 x 5 cards are index files. Also, most abstract/extract/summary files are in 5 x 8 card form, and most 5 x 8 card files contain abstracts/extracts/summaries. Document files are most often in hard copy form versus "other" forms. "Data" files at present are in three forms but mostly in magnetic tape form. Although not shown in Figure III v (77), some 50 million of the some 60 million punched cards in the data base are index records. How much index information must be controlled before it is considered content data is relative to the user and his purpose. Variances between agencies in the form-content correlation factor are small, although the occurrence volumes of a given form-content combination vary considerably between agencies. That is, although content categories and form preferences vary between agencies, the form-content correlations generally hold. Of particular interest to this study are the "index" files, and further analysis of these files was made as discussed in the next section under "problems."

u. Filing Criteria

In addition to the foregoing file parameters, there was recorded during the field survey a brief narrative statement of the qualitative criteria used in selecting items to go into a given file. It is in these statements that subject, area, source, and even more subtle file criteria are reflected. This is qualitative nonformatted information and has not been analyzed as yet. 36/

- 46 -

TOP SECRET

25X1

~~SECRET~~B. Information Processing ProblemsIntroduction

For the past 10 years, most elements of the intelligence community have recognized that an IP problem exists. Characterized symptomatically, the problem appears to be a serious imbalance between the quantity of information that must be collected, processed, and analyzed and the amount and type of resources available to do so. The problem is aggravated by increasing input volumes and severe time constraints.

Much effort has been expended in examining IP concepts, characteristics, and techniques for the valuable insights that might be provided in alleviating the more difficult aspects of the problem. Collectively the community has expended considerable effort in discussion, study, research, experimentation, and further discussion -- provided "business as usual" was not adversely affected. This restriction would seem patently logical and vitally necessary, for these efforts have been carried on in operating environments, each with long-standing commitments to meet. Apparently, ad hoc adjustment usually is substituted for system optimization. Significantly, those making these adjustments have shown great ingenuity in avoiding system breakdown. The value of these efforts should be questioned ultimately if they forestall demonstrably necessary system overhaul.

Elements of the community most often take action in response to immediate operational problems. It is easy to understand an organization's decision to act independently. Implications of problems that are identified "close to home" are much clearer. Orientation to local problem manifestation conveniently limits scope and definition. This orientation, however, conveys the impression that independent solution is proper regardless of problem scope and definition. Frequently, when the total scope of the problem becomes apparent at the local level, other involved components may not be prepared to participate in the solution even if they agree that the problem is a problem and should be solved. On the other hand, if the total scope of the problem is accepted by the involved community elements, hope of expeditious solution in keeping with the demonstrated urgency of the requirement is in immediate jeopardy. In any event, independent action by elements of the community is directed against a backdrop of immediate operational need, which, to date, has been sufficient to justify lack of intersystem compatibility. Unfortunately, such independent actions have rarely resulted in optimum application of critical resources even within the action organization.

In many cases, solution to locally evident problems has been sought through the introduction of data processing equipment. The

- 47 -

~~SECRET~~

equipment is usually employed in the on-going operation rather than as an integral part of a redesigned operation that would exact from the people and equipment their combined potential for solving critical problems in an efficient, effective, and timely manner.

Simultaneously with the local problem/solution approaches discussed above, several large-scale departmental IP research and development efforts have occurred. In those few cases where significant "operations research" and systems analysis have been done prior to establishing performance requirements and specifications for the ensuing detail system design, they have been constrained to departmental factors and in many cases to only certain functional elements within the department. Enthusiastic scrimmaging occurs whenever the designers or proponents of a given "system" consciously or unconsciously attempt to impose design hegemony over other designs independently developed from different, though similarly limited, performance requirements, specifications, or objectives.

This scrimmaging at least indicates that there is some recognition of the existence of critical IP implications that cut across organization lines. However, no vehicle stronger than ad hoc coordination has ever been instituted to identify and reduce unnecessary duplication of effort or to resolve apparent problems on a systematic community-wide basis. In mitigation, it must be acknowledged that past departmental agency studies have already produced a spectrum of problems sufficiently broad in scope and impact to frighten even the system zealot. Expansion to encompass the total environment represented by the intelligence community has never been seriously considered until now.

Whatever success is achieved locally or nationally will be in direct proportion to the ability to isolate and describe the basic information processing context of the community. If second-order and third-order problems can thereby be isolated and adequately described, their solution should result in logical follow-on solution of first-order and fourth-order problems that now seem to occupy center stage in our deliberations. A first-order problem, for example, is, Should installations be considered organizations in the Intelligence Subject Code? In response to a fourth-order problem, for example, it might be concluded that a common centralized information processing and reference system should result in the exchange of processed information without duplicate processing.

An alternative fourth-order conclusion is that a single indexing/coding system should be applied by all participants to allotted portions of all documents. In either case the system design should provide retrieval by document or information item, be subject/content

SECRET

and use-oriented, be both general and discrete in response pattern, be stable yet dynamic, and provide for "browsing" and information/item correlation -- all of which is by way of saying that such a design is an impossibility. The best system that can be offered even conceptually (not counting dreams) for this Nth order problem is a design that satisfies most of these requirements in a general way but none of them in a specific way. The result is partial acceptance (usually for the sake of appearance), mediocre performance, and continued application of specific use-oriented indexing/coding designs unique to each environment.

What is the keystone problem here? Is it that "we have no information theory?" Is it only our inability to design the requisite code? Have we just lacked the authoritative legislation? Or is it that users perform many unique operations on the identical information and that perhaps standardization would impart an undesirable inflexibility?

This section of the report will attempt to address or identify some of the second-order and third-order problems. Even more specific second-order and third-order problems also are addressed in Appendix F and in Appendix H, Volume VI. After viewing the "picture" of the community's present "system" in A, above, it is assumed that the reader is no longer expecting the single simple answer. Also, what follows is not so definitive as it is a sampling of what problem analysis could and should be done on the data presented in A, above, and on data residing still unexploited in SCIPS files as well as data that is still uncollected. Unlike the quantitative picture section, this section, while capitalizing on the data section, also permits the employment of long-suppressed qualitative biases or "experiences."

1. Indexing

There are two aspects to indexing: (1) the indexing codes or system and (2) the resulting physical index, usually in the form of symbols on the item in file or a separate file of index records.

a. Index Files

Of the 926 files (210 million unit records) surveyed in Stage I, 229 files (92 million records) can be considered as physical index files, almost all of them separate from the item indexed. In other words, more than 40 percent of items in files are primarily index records. The distribution of these files by agency is shown in Figures III x (1) and (2). The proportions of unit records between agencies for index files is much the same as for all files surveyed. As can be seen in Figure III x (3), punched cards and 3 x 5 cards are

- 49 -

SECRET

Page Denied

SECRET

2. Data Exchange

a. General

Most member organizations of USIB seem to operate one or more intelligence "systems." That is, they engage in all phases of an operation from directing and arranging acquisition of potentially useful information to the point where the information resides in files ready for recall and further use in conjunction with newly acquired information. These systems are buttressed by collection management processes -- requirements, evaluation, and feedback.

Member organizations also participate in the exploitation of information derived from other organization "systems." The across-"system" aspect associated with the introduction of disparate equipment by USIB members is the basis for an increasing concern for "compatibility" between "systems."

These "systems" appear in the main to orient to their respective methods of acquisition. Thus processing is tied to reconnaissance systems -- photographic and human observations. Likewise, processing is built around intercept systems wherein COMINT, ELINT, and RADINT are established systems [redacted] are systems under development. The exploitation of foreign broadcast intercept maintains its system integrity. We also find system integrity in the acquisition and exploitation of foreign publications.

25X1

A given item of information may be acquired by more than one of these "sources." This may not be readily apparent, because of (1) varying reporting criteria; (2) different transformation requirements inherent in "system" take; (3) different transmission media; and (4) different formatting, indexing, and coding techniques applied by various agencies and components.

The ability to move between these "systems" is a function of access clearance both personal and organizational, knowledge of "system" existence, input and output capabilities, and the time available. In addition, outputs from the "system" usually are made available to other elements of the community on a need-to-know basis.

At present, then, information is exchanged through interagency agreement usually, but not always, in hard copy form. In addition, certain areas of concern have been identified, such as biographic collection, analysis, and reporting, and apportioned among components concerned, who then perform these functions for the community. A third method enhancing information exchange is provided by allowing other agency personnel access to various activities, products, and

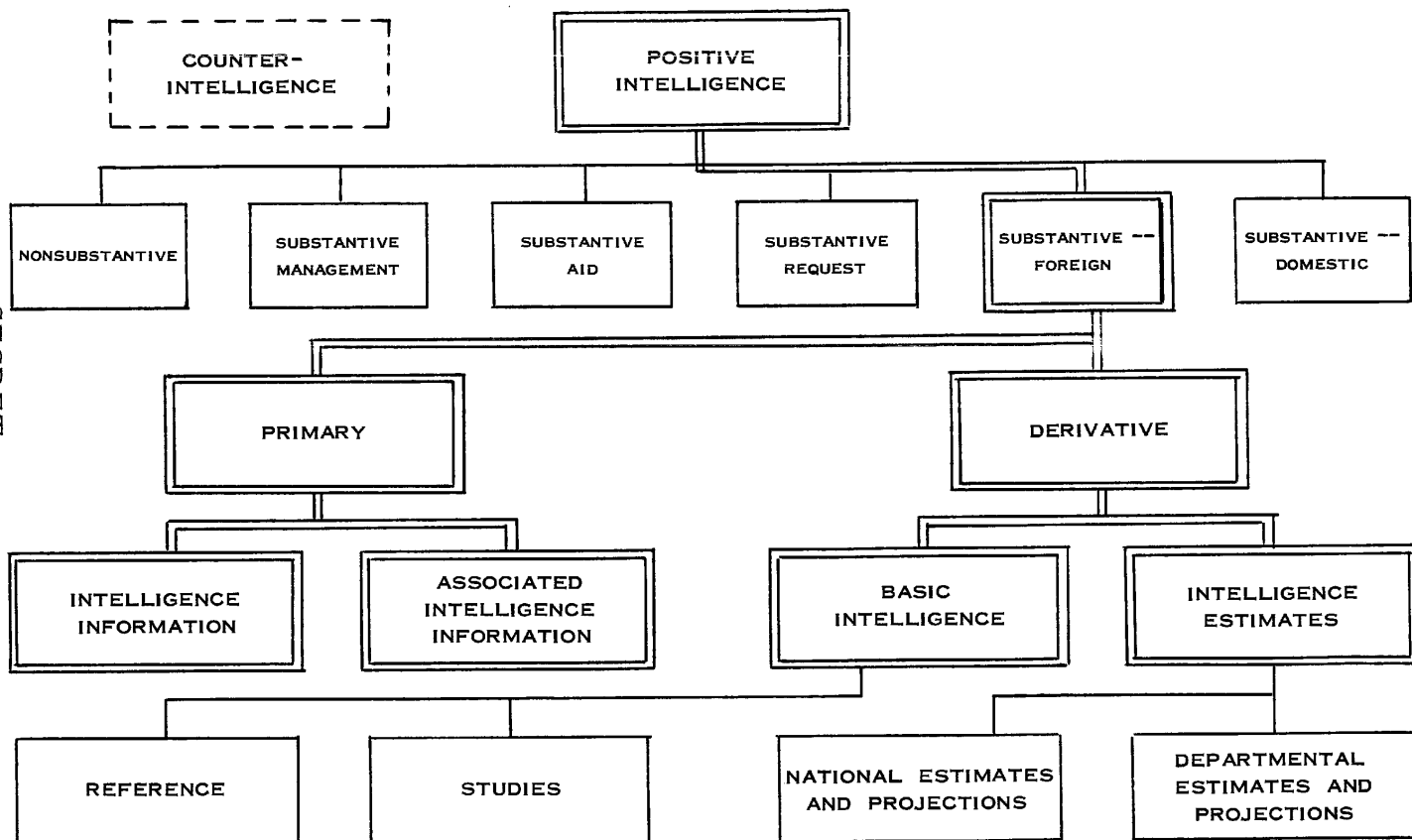
SECRET

25X1

Page Denied

Next 20 Page(s) In Document Denied

CATEGORIZATION OF INTELLIGENCE
AND
INTELLIGENCE INFORMATION



Page Denied

Next 2 Page(s) In Document Denied

SECRET

C. General Discussion

Included in this section are short qualitative discussions on those topics or problem areas that are of current and continuing concern to the intelligence IP community. This section gives some vent to the frustration of not having had the time or resources to do a thorough study and analysis of the problem areas addressed. Redundancy and some inconsistencies will be detected as will an angry tone -- it is hoped that these will not seriously impair the objective of stimulating some creative thought.

1. The Information Flood

Most intelligence organizations disseminate incoming information so that it may be examined by analysts for pertinence to estimates that they are maintaining. Most intelligence organizations find it necessary to file much of the same incoming information so that it may be recalled for retrospective analysis. The initial dissemination to users for continuing analysis and updating of estimates has proved to be quite effective in handling large volumes of incoming information in a timely manner. This process coupled with "working" files (usually analyst-established and analyst-maintained) makes it possible to provide up-to-date estimates in assigned areas of functional responsibility in a relatively short time. A process of developing estimates on request, by retrieving pertinent information from general-purpose central files for subsequent analysis, might involve weeks or even months before comparable answers could be obtained.

On the other hand, the initial dissemination, continuing analysis, and estimate update process does not respond effectively to the solution of newly identified problems or even to unanticipated aspects of old ones. In these cases the need arises to re-examine information that has been acquired and processed over a protracted period of time. Such requirements for retrospective research can be satisfied only by comprehensive storage and retrieval capabilities.

It appears that the two problem solutions -- that is, timely exploitation through continuing analysis and versatile exploitation through retrospective research and analysis -- are much more complementary than competitive. There is a question, however, as to whether the respective capabilities of the two procedures for solving problems are more complementary than competitive in providing for timely exploitation on the one hand and for versatile exploitation on the other.

SECRET

Page Denied

Next 252 Page(s) In Document Denied